Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

# Macroeconomic Factors Strike Back: A Bayesian Change-Point Model of Time-Varying Risk Exposures and Premia in the U.S. Cross-Section

*Daniele Bianchi, Massimo Guidolin and Francesco Ravazzolo*

**Working Paper n. 550**

**This Version: June 6, 2015**

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

# Macroeconomic Factors Strike Back: A Bayesian Change-Point Model of Time-Varying Risk Exposures and Premia in the U.S. Cross-Section*

Daniele Bianchi[†], Massimo Guidolin[‡] and Francesco Ravazzolo[§]

**Abstract**

This paper proposes a Bayesian estimation framework for a typical multi-factor model with time-varying risk exposures to macroeconomic risk factors and corresponding premia to price U.S. publicly traded assets. The model assumes that risk exposures and idiosynchratic volatility follow a break-point latent process, allowing for changes at any point on time but not restricting them to change at all points. The empirical application to 40 years of U.S. data and 23 portfolios shows that the approach yields sensible results compared to previous two-step methods based on naive recursive estimation schemes, as well as a set of alternative model restrictions. A variance decomposition test shows that although most of the predictable variation comes from the market risk premium, a number of additional macroeconomic risks, including real output and inflation shocks, are significantly priced in the cross-section. A Bayes factor analysis massively favors of the proposed change-point model.

**Keywords:** Structural breaks, Stochastic volatility, Multi-factor linear models, Asset Pricing.

**JEL codes:** G11, E44, C11, C53

*This version: June 6, 2015, We are grateful to Carlos M. Carvalho, Fabio Trojani (discussant) and to seminar participants at the University of St.Gallen, Bocconi University, Norges Bank, the $1_{st}$ Vienna workshop on high-dimensional time series in macroeconomics and finance, the $5_{th}$ Italian Congress of Econometrics and Empirical Economics, the SIS 2013 Statistical Conference, the $6_{th}$ International Conference on Computational and Financial Econometrics and the 2014 European Financial Association meetings for helpful comments and suggestions. Ravazzolo is also affiliated with the BI Norwegian Business School. The views expressed in this paper are our own and do not necessarily reflect those of Norges Bank.

†University of Warwick, Warwick Business School, Coventry, CV4 7AL, UK. Daniele.Bianchi@wbs.ac.uk

‡Department of Finance and IGIER, Bocconi University, Milan, Italy. massimo.guidolin@unibocconi.it.

§Norges Bank and BI Norwegian Business School, Oslo, Norway. Francesco.Ravazzolo@Norges-Bank.no.

# 1 Introduction

Can a selected set of macroeconomic variables explain the cross-sectional behavior of U.S. stock and bond returns, i.e., why different assets earn different average rates of return? This simple question lies at the heart of the burgeoning field of macro-finance. Remarkably enough, the answer provided by at least 20 years of research on this crucial question has been predominantly negative (see e.g., Chan et al. 1998; McCulloch and Roley 1993; Shanken and Weinstein 2006): although occasional nuances to this fundamentally negative result have been reported (e.g., Flannery and Protopapadakis 2002; Kramer 1994), it is common wisdom that macroeconomic factors can hardly explain the cross-sectional dynamics of asset valuations and returns of U.S. stock and bond portfolios. Such a disconnect between changes in aggregate variables representing sources of systematic risk—like in the case of output and inflation growth news—and asset returns has long represented a puzzle.

In this paper we propose and estimate through Bayesian methods a flexible parametric multi-factor, stochastic volatility asset pricing model in which both risk exposures (betas) and the prices of a number of macroeconomic risk factors are time-varying and effectively explain the cross-section of U.S. stock and bond returns (see Gungor and Luger 2013). Time variation is modelled as a latent, change-point process. We show that an explicit parameterization of latent change-points in betas and risk premia plays a dominant role. By comparing our baseline model with restricted versions of the same, we also provide evidence that both stochastic volatility and infrequent but possibly large parameter instability are key drivers of the capability of the model to capture cross-sectional return dynamics.

Drawing a precisely estimated link between time-varying betas on selected macroeconomic risk factors and stock and bond excess returns also speaks to the very heart of finance theory, because any evidence uncovered bears on the fundamental issue of the key features of the general pricing mechanism, called the stochastic discount factor (SDF), underlying observed security prices. Practically, the SDF depends on the shape of the (aggregate) risk aversion function of investors and therefore reflects the way in which systematic risk factors are priced in the aggregate (see e.g., Cochrane 2001; Singleton 2006). In our paper we show that is both possible and useful to connect such an SDF (assumed to exist and to be unique) and macroeconomic risks.

A related question concerns the most appropriate methods available to researchers to learn about such the SDF. Our paper offers a contribution to an extensive literature on the estimation of empirical SDFs, specializing to a particular set of linear multi-factor models, offers a novel statistical framework to implement such models, and shows how this works using an empirically relevant application. With reference to an application to 40 years of monthly data on excess returns on 23 key portfolios of securities traded in the U.S., we show that while commonly used methods to estimate macro-based linear factor models fail to lead to sensible conclusions, an encompassing Bayesian estimation scheme that allows for both parameter uncertainty and instability in factor exposures and risk premia delivers encouraging results.

Following the seminal work of Fama and MacBeth (1973), two-step multi-factor asset pricing models (MFAPMs) have been commonly used to estimate multi-factor models. Fama-MacBeth's (henceforth F-MB) approach, first proposed for the plain vanilla CAPM but then extended to a wider class of linear models, corresponds to a very simple algorithm: the risk premium on any asset or portfolio is decomposed as the sum of risk exposures to a number of risk factors multiplied by the associated unit price for each factor. The algorithm uses a first set of rolling window, time series regressions to obtain estimates of the betas, followed by a second-pass set of cross-sectional (across assets) regressions that using the first-pass risk exposures as inputs to derive time-varying estimates of the premia. The limitations of this methodology are now well-understood: most inferential statements made as a result of the second-pass would be valid if and only if one could assume that the first-pass betas were fixed in repeated samples, which contradicts their random nature deriving from their being least squares estimates. Unless additional assumptions are introduced, this creates a problem with generated regressors being used in the second-step, which makes most of the inferential statements commonly made when the resulting error-in-variables problems are ignored invalid (see Pagan 1984). F-MB's approach also suffers from another problem: although identifying time-variation in risk exposures and premia with a rolling window least square estimation is robust because it is nonparametric, the length of the window is usually chosen in an arbitrary way and this can result in a severe loss of efficiency (see e.g., Maheu and McCurdy 2009).

To overcome these problems, we introduce a different approach where time variation in risk exposures and premia is explicitly modelled as a change-point process. Specifically, we model risk exposures as latent stochastic processes in a mixture innovation framework as in

Giordani and van Dijk (2007), Giordani and Kohn (2008), Groen et al. (2013), Maheu and Gordon (2008). The parameters of interest are constant unless a break-point variable takes a unit value, in which case the parameters are allowed to jump to a new level, as a result of a normally distributed shock. Furthermore, to consistently overcome the problems with generated regressors, the model is estimated in a single step by using a Bayesian approach, following the seminal work by McCulloch and Rossi (1991) and Geweke and Zhou (1996). By construction, our approach provide a single-step procedure that yields exact inferences on MFAPMs. Moreover, our approach makes it possible to compute the posterior distribution of virtually any function of the parameters that can be useful to implement economic tests.

Our main results can be summarized as follows. First, using a variety of metrics—such as Bayes factors and average pricing error performance—we obtain evidence of the importance of capturing both instability in betas and in stochastic volatility; additionally, simpler time-varying parameter models in which betas follow random walk processes appear to be outperformed by our change-point model. Moreover, a variance decomposition test shows that by considering model instability, together with parameter uncertainty, the amount of cross-sectional excess return variation explained by the factor model increases with respect not only to a standard F-MB, but also with respect to the case in which specific parsimonious restrictions on the model dynamics are imposed. Second, the Bayesian time-varying betas, stochastic volatility model leads to economically realistic estimates with reference to an application for which the standard two-stage approach fails to provide plausible insights and would lead to a MFAPM rejection. For instance, a two-step F-MB approach leads to display a positive and statistically significant cross-sectional pricing error. On the contrary, in the Bayesian case, the average value of the posterior medians of the same parameter often indicate the absence of systematic mis-pricing. Third, we show that idiosyncratic risks play a significant role, and peaks around the early 2000s and the end of our sample.

The remainder of the paper is organized as follows. Section 2 outlines the framework for our MFAPM, which includes a detailed discussion of the underlying Bayesian methodology. Section 3 describes the dataset and how we operationalize our model by determining prior hyper-parameters. Section 4 shows the main empirical results. Next, in Section 5 we provide a further economic assessment of the model performance. Section 6 concludes.

## 2 A Bayesian Framework for Linear MFAPMs

Our empirical work is based on a model from the multi-factor linear class introduced by Ferson and Harvey (1991). Multi-factor asset pricing models (MFAPMs) posit a linear relationship between asset returns and a set of (macroeconomic, systematic) factors that are assumed to capture business cycle effects on beliefs and/or preferences (as summarized by a SDF with time-varying properties, see e.g., Cochrane 2001) and hence on risk premia. The advantage of MFAPMs consists of the fact that a number of systematic risk factors $K << N$ may efficiently capture relatively large portions of the variability in the cross-section of returns. If we call the process for the risk factors $F_{j,t}$ $(j = 1, ..., K)$ and $r_{i,t}$ the period *excess* return on asset or portfolio $i = 1, ..., N$, computed as $r_{i,t} \equiv [(P_{i,t} - P_{i,t-1} + D_{i,t})/P_{i,t-1}] - r_t^f$ where $P_{i,t}$ denotes the price of any asset or portfolio, $D_{i,t}$ any dividend or cash flow paid out by the asset, and $r_t^f$ the one-period interest rate, a typical MFAPM can be written as:

$$r_{i,t} = \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \epsilon_{i,t} \qquad \epsilon_{i,t} \sim N(0, \sigma_{i,t}^2), \tag{1}$$

where $E[\epsilon_{i,t}] = E[\epsilon_{i,t} F_{j,t}] = 0$ for all $i = 1, ..., N$ and $j = 1, ..., K$. The time-varying processes for risk factors exposures, $\beta_{ij,t}$, and idiosyncratic risk, $\sigma_{i,t}^2$, are left unspecified by asset pricing theory. Standard approaches in empirical finance posit that betas depend on a pre-specified set of instruments, which can capture either macroeconomic or firm-specific news on single assets risk exposures (see e.g. Lettau and Ludvigson 2001 and Nardari and Scruggs 2007). However, estimates of betas obtained using instrumental variables are by construction very sensitive to the choice of instruments (e.g. Harvey 2001). Simple GARCH(1,1) or stochastic volatility represent instead the benchmark to model time variation in idiosyncratic risk. Finally, the $\beta_{i0,t}$ coefficients are often interpreted as abnormal returns on asset $i$ "left on the table" after all risks ($F_{j,t}$, $j = 1, ..., K$) and risk exposures ($\beta_{ij,t}$, $j = 1, ..., K$) have been taken into account.

In the Merton (1973) inter-temporal CAPM (ICAPM), the returns generating process in (1) implies that the SDF, $M_{t+1}$, must be linearly dependent on the $K$-dimensional vector of macroeconomic risk factors $M_{t+1} = a_t + b_t' F_{t+1}$. Let $\tilde{F}_t = (1, F_t)$, under no-arbitrage opportunities, the fundamental pricing equation $E_t\left[M_{t+1} \cdot r_{i,t+1}\right] = 0$ implies that

$$E_t[r_{i,t+1}] = \frac{Cov_t\left(F_{t+1}, r_{i,t+1}\right)}{Cov_t\left(F_{t+1}, F'_{t+1}\right)} \times -r_t^f Cov_t\left(F_{t+1}, \tilde{F}_{t+1}\right) b_t = \beta'_{i,t}\lambda_t \qquad (2)$$

namely, the expected excess return on asset $i$ over the interval $[t, t+1]$, $E_t[r_{i,t+1}]$, are related to its "current" betas, $\beta_{i,t} = [\beta_{i1,t}, \beta_{i2,t}, ..., \beta_{iK,t}]'$, and the factors risk premia, $\lambda'_t = [\lambda_{1,t}, \lambda_{2,t}, ..., \lambda_{K,t}]$ (see Cochrane 2001 for more details). By construction, both the betas and the risk premia are conditional on the information publicly available at time $t$, that capture any effects of the state of the economy on unit risk premia (see e.g., Bossaerts and Green 1989). The framework in (1)-(2) describes a general conditional equilibrium asset pricing framework that is known to hold under a variety of alternative assumptions.

The standard approach to test the equilibrium asset pricing model shown in (1)-(2), is the two-stage procedure *à* la Fama and MacBeth (1973). In the first stage, for each of the assets, the factor betas are estimated using time-series regressions from historical excess returns on the assets and economic factors. That is, for month $t$, (1) is estimated using the previous, say, sixty months (ranging from $t-61$ to $t-1$) in order to obtain estimates for the betas, $\hat{\beta}^{60}_{ij,t}$. This time-series regression is updated each month. The choice of a 60-month rolling window scheme is typical of the literature. In the second stage, the equilibrium restriction (2) is estimated for each of the periods in our sample a cross-sectional regression using ex-post realized excess returns:

$$r_{i,t} = \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t}\hat{\beta}^{60}_{ij,t} + \zeta_{i,t} \qquad \zeta_{i,t} \sim N(0, \tau) \qquad (3)$$

for $i = 1, ..., N, t = 61, ..., T$. The equilibrium condition (2) makes clear that $\lambda_{0,t}$ should equal zero if the model is correctly specified. Indeed, in the absence of arbitrage all zero-beta assets should command a rate of return that equals the short-term rate, which is null in our case. Therefore, cross-sectional tests of multi-factor models boil down to evaluate the importance of the economic risk variables by evaluating whether their risk premiums are priced and whether, on average, the coefficients $\hat{\lambda}_{0,t}$ are not significantly different from zero. Clearly, this $T$ cross-sectional regressions simply test (2) in a nonparametric fashion, in the sense that any resulting time variation in the $\lambda_{0,t}$ and $\lambda_{j,t}$ coefficients fails to be explicitly and parametrically related to any of the instruments assumed by the researcher.

Although widely used, the two-stage Fama-MacBeth (henceforth F-MB) approach has a number of statistical drawbacks (see Petersen 2009 for a complete discussion). First, the second stage multivariate regression used to test for the equilibrium restriction (2) suffers from obvious generated regressor (error-in-measurement) problems as the estimated first-stage, rolling window beta estimates $\hat{\beta}^{60}_{ij,t-1}$ are used as regressors on the right-hand side. For instance, Ang and Chen (2007) have stressed that when the cross-sectional estimates of the betas $\hat{\beta}^{60}_{ij,t-1}$ co-vary with the underlying but unknown risk premia, (3) may easily yield biased and inconsistent estimates of the risk premia themselves. Unfortunately, this co-variation is extremely likely: for instance, the asset pricing literature generally presumes that during business cycle downturns both the quantity of risk (the size of the betas) and the unit risk prices would increase, simply because recessions are characterized by higher systematic uncertainty as well as by lower "risk appetite". Second, for instance as emphasized by Jostova and Philipov (2005) with reference to a single-factor conditional CAPM, when parameters in linear asset pricing models are estimated from the data, their uncertainties should be taken into account. Third, the need to perform the estimation of (1)-(2) in two distinct stages that use rolling windows to capture parameter instability is not only *ad hoc* but also inefficient because the lack of more specific parametric forms makes testing for time-variation very hard and dependent on hard-to-justify choices of the rolling window length, the updating rules applied to select whether constant or decaying weights should be applied, etc. (see Maheu and McCurdy 2009).

## 2.1 Our Change-Point Model

The previous discussion of the standard F-MB two-step procedure implies that for testing the equilibrium asset pricing model shown above, we need to: (1) avoid using estimates of the first-stage betas as if these were observed variables that may be constant in repeated samples; (2) fully account for parameter uncertainty; and (3) sensibly capture parametric instability to reflect the commonly perceived (and tested) fact that both the relationship between excess returns and factors, namely risk exposures ($\beta_{ij,t}$), the risk premia ($\lambda_j$, for $i = 1, ..., N$ and $j = 1, ..., K$), and possibly also residual idiosyncratic variances ($\sigma^2_{i,t}$) change over time.

We therefore develop a new Bayesian estimation approach in which: (1) the error in measurement is avoided following McCulloch and Rossi (1991) and Geweke and Zhou (1996), by characterizing the joint posterior distribution of the risk exposures and the risk premia such that both states and parameters are jointly estimated in a single step; (2) parameter uncertainty is fully addressed by using Bayesian techniques; and (3) model instability is captured by introducing stochastic breaks in the dynamics of the factor loadings as well as of idiosyncratic risks.

Specifically, we characterize the relationship between excess returns, factors and risk premia, as well as the time-varying dynamics in factor loadings and idiosyncratic volatility in a state-space form where we jointly consider the linear factor model (1) (i.e. observation equation) and the non-linear no-arbitrage restriction (2);

$$r_{i,t} = \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \sigma_{it} \epsilon_{i,t} \qquad \epsilon_{i,t} \sim N(0,1) \tag{4}$$

$$r_{i,t} = \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t} \beta_{ij,t-1} + e_{i,t} \qquad e_{i,t} \sim N(0,\tau^2) \tag{5}$$

where $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t}, ..., \epsilon_{N,t})' \sim N(0, I_N)$ and $E[\epsilon_{i,t}] = E[\epsilon_{i,t} F_{j,t}] = E\left[e_{i,t} \beta_{ij,t-1}\right] = 0$ for all $i = 1, ..., N$ and $j = 1, ..., K$. The error term $e_{i,t}$ is due to the fact that (5) just represents a statistical approximation of the equilibrium condition (2). Indeed, such no-arbitrage restriction holds perfectly only if the number of assets/portfolios goes to infinity. In finite samples, and with a finite number $N$ of assets, this residual term can be arbitrarily large. The time varying parameters $\beta_{ij,t}$ and $\sigma_{it}$ are described by the state equations

$$\beta_{ij,t} = \beta_{ij,t-1} + \kappa_{ij,t} \eta_{ij,t} \qquad j = 0, ..., K, \tag{6}$$

$$\ln(\sigma^2_{i,t}) = \ln(\sigma^2_{i,t-1}) + \kappa_{iv,t} \upsilon_{i,t} \qquad i = 1, ..., N, \tag{7}$$

where $\eta_{i,t} = (\eta_{i0,t}, \eta_{i1,t}, ..., \eta_{iK,t}, \upsilon_{i,t})' \sim N(0, q_i^2)$ with $q_i^2 = diag(q_{i0}^2, q_{i1}^2, ..., q_{iK}^2, q_{iv}^2)$. Stochastic variations (breaks) in the level of both the beta coefficients and of the idiosyncratic variance $\sigma^2_{it}$ are introduced and modeled through a mixture innovation approach as in Ravazzolo et al. (2007) and Giordani and Kohn (2008). The latent binary random variables $\kappa_{1ij,t}$ and $\kappa_{2i,t}$ are used to capture the presence of random shifts in betas and/or idiosyncratic variance (see Mitchell and Beauchamp 1988; George and McCulloch 1993; Miazhynskaia et al. 2006). The random variable $\kappa_{1ij,t}$ takes then a value equal to one if a structural break for the $j$th factor in the equation for the $i$th asset at time $t$ takes place. We assume that the structural breaks are independent of each another (i.e., across assets as well as factors) and over time, with:

$$\Pr\left[\kappa_{ij,t}=1\right]=\pi_{ij} \qquad \Pr\left[\kappa_{iv,t}=1\right]=\pi_{iv} \qquad i=1,...,N \quad j=0,...,K \tag{8}$$

This specification is very flexible as both risk exposures, $\beta_{ij,t}$, and idiosyncratic risks, $\sigma^2_{i,t}$, are allowed to change on every time period, but they are not imposed to be changing at every point in time. In our view, this helps to side-step the difficult (if not impossible) task of persuading a Reader that the assumed dynamics represents the "right" kind: given our uninformative priors, if the data need frequent breaks in betas of a small size, the posterior of the corresponding parameters will provide indications in this direction; similarly, if the data need a (set of independent) stochastic volatility process(es) with frequent shifts in idiosyncratic variance, posterior estimates will give appropriate indications, etc. Note that we can interpret $q_i$ as the "size" of the break: a large $q_{ij}$ means for instance that whenever $\beta_{ij,t}$ is hit by a break, i.e. $\kappa_{ij,t}=1$, such a shift is more likely to be large (in absolute value).

The model presented in (4)-(8) is the most general specification we consider in this paper. We will call this model B-TVB-SV indicating that we consider a Bayesian (B), Time-Varying Betas (TVB) and Stochastic Volatility (SV) framework. Here the words time-varying and stochastic for the betas and the volatility are synonymous of structural breaks. For comparative purposes, we consider a two alternative restrictions on the model dynamics: (1) the restriction $\kappa_{iv,t}=0$ $\forall i$, $t$ in (6) in which case idiosyncratic risk is assumed to be constant. We will call this model a Bayesian (B) homoskedastic Time-Varying Betas (TVB) model, i.e. B-TVB; and (2) the restriction $\kappa_{ij,t}=1$ $\forall i$, $j$, $t$ and $\kappa_{iv,t}=1$ $\forall i,t$, in (6)-(7), in which case the parameters follow a random walk and structural breaks occur at each time $t$ (see e.g., Koop and Potter 2007; West and Harrison 1997). We call this Bayesian (B) Time-Varying Parameters (TVP) model, i.e. B-TVP. The constant volatility B-TVB and the random walk B-TVP specifications are useful to highlight the effects of instabilities in residual variances and to show the benefit of considering more parsimonious, occasional breaks in (6)-(7) as opposed to frequent (continuous) breaks (see Giordani and Villani 2010, for a related discussion).

### 2.2 Prior Specification

For parameter inference in (5)-(7), we use a Bayesian approach. Such approach allows to incorporate parameter uncertainty in testing the MFAPM in a natural way. Also, Bayesian inference allows to characterize the posterior distribution of virtually any function of the model parameters. For instance, we can characterize the posterior distribution of $\kappa_{ij,t}$ and $\kappa_{iv,t}$ for $i = 1, ..., N$, $j = 1, ..., K$ and $t = 1, ..., T$, which can be used to incorporate uncertainty on the timing of structural breaks.

For each of the *ith* asset/portfolio, the parameters of the model (4)-(8), are the structural break probabilities $\pi_i = (\pi_{i0}, \pi_{i1}, ..., \pi_{iK}, \pi_{iv})'$, and the vector of the size of the breaks $q_i^2 = \left(q_{i0}^2, q_{i1}^2, ..., q_{iK}^2, q_{iv}^2\right)'$. We collect the model parameters in a $(2K+2)$-dimensional vector $\theta_i = \left(\pi_i', q_i^{2'}\right)'$. For the Bayesian algorithm to work, we need to specify the prior distributions for each of the parameters. For the structural break probabilities, we take Beta distributions

$$\pi_{ij} \sim Beta(a_{ij}, b_{ij}) \quad \pi_{iv} \sim Beta(a_{iv}, b_{iv}) \qquad \text{for} \quad i = 1, ...N, \quad j = 1, ...K. \tag{9}$$

The parameters $a_{ij}, b_{ij}$ and $a_{iv}, b_{iv}$ represent the shape hyper-parameters and can be set according to our prior beliefs about the occurrence of structural breaks. The expected prior probability of a break in $\beta_{ij,t}$ and $\ln(\sigma_{i,t}^2)$ is given by $a_{ij}/\left(a_{ij} + b_{ij}\right)$ and $a_{iv}/\left(a_{iv} + b_{iv}\right)$, respectively. For the conditional variance parameters, which reflect our prior beliefs about the size of the structural breaks, we assume an inverted Gamma-2 prior,

$$q_{ij}^2 \sim IG - 2(\overline{\gamma}_{ij}, \delta_{ij}) \qquad q_{iv}^2 \sim IG - 2(\overline{\gamma}_{iv}, \delta_{iv}) \quad \text{for} \quad i = 1, ...N, \quad j = 1, ...K \tag{10}$$

where $\overline{\gamma}_{ij} = \gamma_{ij}\delta_{ij}$ and $\overline{\gamma}_{iv} = \gamma_{iv}\delta_{iv}$. The expected prior break size for the betas (log-volatility) equals the square root of $\gamma_{ij}\delta_{ij}/\left(\delta_{ij} - 2\right)$ for $\delta_{ij} > 2$ ($\gamma_{iv}\delta_{iv}/\left(\delta_{iv} - 2\right)$ for $\delta_{iv} > 2$). The density for the joint prior $p(\theta_i)$ is given by the product of the prior specifications (9)-(10). Note the priors are independent across assets/portfolios.

The set of risk premia $\lambda_t$ in (5) is not dynamic in nature. In fact, time variation is inherited from the dynamics of the exposures to macroeconomic risk factors $\beta_{ij,t-1}$. This simplify the estimate of the no-arbitrage restriction as a simple multi-variate linear regression. We take at each time $t$ an independent prior structure of the form

$$\lambda \sim MN(\underline{\lambda}, \underline{V}) \qquad \tau^2 \sim IG-2(\overline{\psi}_0, \Psi_0) \tag{11}$$

The parameters $\underline{\lambda}$ and $\underline{V}$ represent the $K \times 1$ location vector and the $K \times K$ scale matrix, while $\overline{\psi}_0 = \psi_0 \Psi_0$ such that the expected prior $\tau^2$ equals $\psi_0 \Psi_0 / (\Psi_0 - 2)$.

## 2.3 Posterior Simulation

Posterior results are obtained through the Gibbs sampler algorithm developed in Geman and Geman (1984) in combination with the data augmentation technique by Tanner and Wong (1987) and Frühwirth-Schnatter (1994). The latent variables $B = \{\beta_{i,t}\}_{i=1}^{N}{}_{t=1}^{T}$, with $\beta_{i,t} = (\beta_{i0,t}, \beta_{i1,t}, ..., \beta_{iK,t})$, $\Sigma = \{\sigma_{it}^2\}_{i=1}^{N}{}_{t=1}^{T}$, $\mathcal{K}_\sigma = \{\kappa_{iv,t}\}_{i=1}^{N}{}_{t=1}^{T}$, and $\mathcal{K}_\beta = \{\kappa_{i,t}\}_{i=1}^{N}{}_{t=1}^{T}$, with $\kappa_{i,t} = (\kappa_{i0,t}, \kappa_{i1,t}, ..., \kappa_{iK,t})$ are simulated alongside the model parameters $\theta = \{\theta_i\}_{i=1}^{N}$ and the risk premia $\lambda = \{\lambda_t\}_{t=1}^{T}$, with $\lambda_t = (\lambda_{0,t}, \lambda_{1,t}, ..., \lambda_{K,t})$. To apply the Gibbs sampler we need to write down the complete likelihood function, namely, the joint density of the data and the state variables.

$$\begin{aligned} & p(R, B, \mathcal{K}, \Sigma | \lambda, \theta, F) \\ & = \prod_{t=1}^{T} \left( \prod_{i=1}^{N} p(r_{it} | \lambda_t, F_t, \beta_{it}, \beta_{it-1}, \sigma_{it}^2) p(\sigma_{it}^2 | \sigma_{it-1}^2, \kappa_{iv,t}, q_{iv}^2) \pi_{iv}^{\kappa_{ivt}} (1 - \pi_{iv})^{1-\kappa_{ivt}} \times \right. \\ & \left. \times \left( \prod_{j=0}^{K} p(\beta_{ij,t} | \beta_{ij,t-1}, \kappa_{ij,t}, q_{ij}^2) \times \pi_{ij}^{\kappa_{ijt}} (1 - \pi_{ij})^{1-\kappa_{ijt}} \right) \right), \end{aligned} \tag{12}$$

where $\mathcal{K} = (\mathcal{K}_\beta, \mathcal{K}_\sigma)$ and $F = \{F_t\}_{t=1}^{T}$, with $F_t = (F_{1,t}, F_{2,t}, ..., F_{K,t})'$ the factors, and $R = \{R_t\}_{t=1}^{T}$, with $R_t = (r_{1,t}, r_{2,t}, ..., r_{N,t})'$ the excess returns. We specifiy the densities that make up (12) in Appendix A. The conditional likelihood $p(r_{it} | \lambda_t, F_t, \beta_{it}, \beta_{it-1}, \sigma_{it}^2)$ can be characterized by combining the no-arbitrage restriction (5) in the observation equation (4). Indeed, for conditional zero mean factors (i.e. $E_{t-1}[F_t] = 0$), the system defined by (4)-(5) implies that $\beta_{i0,t} \simeq \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t} \beta_{ij,t-1}$. By plugging such non-linear restriction in (4) we can obtain the conditional likelihood. Combining the prior $p(\theta)$ with (12), we obtain the posterior density $p(\theta, B, \mathcal{K}, \Sigma, \lambda | R, F) \propto p(\theta, \lambda) p(R, B, \mathcal{K}, \Sigma | \lambda, \theta, F)$. Our Gibbs sampler is a combination of the Forward Filtering Backward Sampling of Carter and Kohn (1994) and Kim et al. (1998), and the efficient sampling algorithm for the random breaks proposed in Gerlach et al. (2000). At each iteration of the sampler we sequentially cycle through the following steps:

1. Draw $\mathcal{K}_\beta$ conditional on $\Sigma, \mathcal{K}_\sigma, \theta$, $R$ and $F$.
2. Draw $B$ conditional on $\Sigma, \mathcal{K}, \theta$, $R$ and $F$.
3. Draw $\mathcal{K}_\sigma$ conditional on $B, \mathcal{K}_\beta, \theta$, $R$ and $F$.
4. Draw $\Sigma$ conditional on $B, \mathcal{K}, \theta$, $R$ and $F$.
5. Draw $\lambda$ conditional on $B$ and $R$.
6. Draw $\theta$ conditional on $B, \mathcal{K}$, $R$ and $F$.

A more detailed description if this Gibbs sampling and the corresponding convergence properties are given in Appendix A and B.

# 3 Data, Prior Choices and Convergence Results

In this section, we outline how to operationalize the B-TVB-SV model on our dataset for the U.S. cross-section of financial returns. In Section 3.1, we discuss this data, whereas in Section 3.2, we discuss our choices for prior hyper-parameters, the sensitivity of posterior estimates to such choices and the convergence results of the algorithm.

## 3.1 Data

We consider a typical application in the empirical finance literature based on a moderate number (23) of monthly time series sampled over the period 1972:01 - 2011:12. The starting date is due to the availability of the complete set of instruments and corporate bond return data. The series belong to two main categories. The first group, "Portfolio Returns" , includes stocks, U.S. Treasuries and notes, and corporate bonds, all organized in portfolios to tame the non-diversifiable risk reflected by excess returns. The stocks are publicly traded firms listed on the NYSE, AMEX and Nasdaq (from CRSP) and sorted according to two criteria. First, 10 industry portfolios are obtained by sorting firms according to their four-digit SIC code. Second, 10 additional portfolios are derived by sorting (at the end of every year, and recursively updating this sorting every year) NYSE, AMEX and Nasdaq stocks according to their size, as measured by the aggregate market value of the company's equity. Using industry and size-sorting criteria to form portfolios of stocks to trade-off "spread" and reduction of idiosyncratic risk, is typical in the literature (see e.g., Dittmar 2002). Moreover, industry- and size-sorting criteria are sufficiently unrelated to make it plausible that industry- and size-sorted equity portfolios may contain non-overlapping information on the underlying factors and risk premia. Data on long- (10-year) and medium-term (5-year) government bond returns are from Ibbotson and available from CRSP. Data on 1-month T-bill, 10-year and 5-year government bond yields and returns are from FREDII at the Federal Reserve Bank of St. Louis and from CRSP. Data on "junk" bond returns are approximated from Moody's (10-to-20 year maturity) Baa average corporate bond yields and converted into return data using Shiller (1979) approximation formula.

The second group collects variables that measure macroeconomic risks. These factors are used as proxies for the systematic, economy-wide forces potentially priced in asset returns. We employ nine factors: the excess return on a wide, value-weighted market portfolio (MKT) that includes all stocks traded on the NYSE, AMEX, and Nasdaq (from CRSP); changes in the default risk premium (DEF) measured as the difference between Baa Moody's *yields* and yields on 10-year government bonds; the change in the term premium (TERM), the difference between 10-year and 1-month Treasury yields; the unexpected inflation rate (UI), computed as the residual of a simple ARMA(1,1) model applied to (seasonally adjusted) CPI inflation rate; the rate of growth of (seasonally adjusted) industrial production (IP); the rate of growth of (seasonally adjusted) real personal consumption (PC); the 1-month real T-bill return computed as the difference between the 1-month T-bill nominal return and realized CPI inflation rate (not seasonally adjusted); the traded Liquidity factor (LIQ) from Pastor and Stambaugh (2003); the Bond premium factor (BPF) from Cochrane and Piazzesi (2005). Using a relatively large number of pre-selected factors is typical of the literature (e.g. Burmeister and McElroy 1988; Chen et al. 1986). Table 1 reports a detailed set of summary statistics.

#### 3.1.1 Traded vs. Non-Traded Factors

One problem with (1) is the difficulty of interpreting $\beta_{i0,t}$ (often called the "Jensen's alpha") when some of the risk factors are not traded portfolios. In principle, any $E[\beta_{i0,t}] \neq 0$ is referred to as an "abnormal" (average) return. However, unless all the factors are themselves tradable portfolios it is impossible to interpret any non-zero $\beta_{i0}$ as an abnormal return (see Gungor and Luger 2013). A factor is tradable if its realizations may be closely replicated ("mimicked" , with a high coefficient of determination) by linear combinations (portfolios) of the test assets

Table 1: Descriptive Statistics

This table reports the descriptive statistics for each of the 23 portfolios used in the empirical analysis as well as the risk factors and the instrumental variables. Data are monthly and cover the sample period 1972:01 - 2011:12.

| **Portfolio/Factor** | **Mean** | **Median** | **Std. Dev.** | **Sharpe Ratio** |
|---|---|---|---|---|
| | **10 Industry Portfolios, Value-Weighted** | | | |
| **Non-Durable Goods** | 1.107 | 1.135 | 4.473 | 0.248 |
| **Durable Goods** | 0.809 | 0.815 | 6.649 | 0.122 |
| **Manufacturing** | 0.988 | 1.195 | 5.195 | 0.190 |
| **Energy** | 1.163 | 0.990 | 5.672 | 0.205 |
| **High-Tech** | 0.924 | 0.950 | 6.897 | 0.134 |
| **Telecommunications** | 0.948 | 1.175 | 4.891 | 0.194 |
| **Shops and Retail** | 0.974 | 1.060 | 5.447 | 0.179 |
| **Healthcare** | 0.990 | 1.050 | 5.094 | 0.194 |
| **Utilities** | 0.933 | 0.995 | 4.142 | 0.225 |
| **Other** | 0.871 | 1.320 | 5.439 | 0.160 |
| | **10 Size-Sorted Portfolios, Value-Weighted** | | | |
| **Decile 1** | 1.073 | 1.205 | 6.347 | 0.169 |
| **Decile 2** | 1.083 | 1.390 | 6.491 | 0.167 |
| **Decile 3** | 1.125 | 1.545 | 6.162 | 0.182 |
| **Decile 4** | 1.089 | 1.500 | 5.952 | 0.183 |
| **Decile 5** | 1.127 | 1.680 | 5.811 | 0.194 |
| **Decile 6** | 1.081 | 1.180 | 5.412 | 0.200 |
| **Decile 7** | 1.088 | 1.255 | 5.382 | 0.202 |
| **Decile 8** | 1.024 | 1.275 | 5.262 | 0.195 |
| **Decile 9** | 0.986 | 1.335 | 4.853 | 0.203 |
| **Decile 10** | 0.844 | 1.075 | 4.473 | 0.189 |
| | **Bond Returns** | | | |
| **10-Year T-Note** | 0.679 | 0.628 | 2.299 | 0.295 |
| **5-Year T-Note** | 0.635 | 0.585 | 1.629 | 0.390 |
| **Baa Corp. Bond (10-20 years)** | 0.831 | 0.863 | 3.237 | 0.257 |
| | **Economic Risk Factors** | | | |
| **Excess Value-Weighted Mkt** | 0.452 | 0.800 | 4.681 | 0.097 |
| **Default Premium** | 0.192 | 0.461 | 3.481 | |
| **Term Spread** | 0.000 | 0.000 | 0.406 | |
| **Industrial Prod. Growth** | 0.186 | 0.256 | 0.755 | |
| **Real Per-capita Cons. Growth** | 0.255 | 0.262 | 0.338 | |
| **Real T-Bill Interest Rate** | 0.087 | 0.102 | 0.357 | |
| **Unexpected Inflation** | 0.000 | -0.016 | 0.301 | |
| **Bond Risk Factor** | 1.093 | 0.982 | 1.944 | 0.562 |
| **Liquidity Factor** | 0.497 | 0.232 | 3.621 | 0.137 |
| | **Instrumental Variables** | | | |
| **Term Yield Spread** | 1.715 | 1.910 | 1.329 | |
| **Credit Yield Spread** | 1.111 | 0.960 | 0.488 | |
| **Dividend Yield** | 3.029 | 2.952 | 1.259 | |

employed in the analysis. Unless all factors are replicated and replaced by the returns on traded portfolios, there may be a considerable difference between the theoretical alphas from an estimated model, and the actual alpha that an investor may harvest from by trading assets on the basis of a MFAPM.

To eliminate such a possibility, we follow the literature (see e.g., Lamont 2001) and proceed as follows. When an economic risk factor is already measured in the form of a return (e.g., this is the case of the U.S. market portfolio, real T-bill rates, the liquidity and bond risk factors, term structure spreads, and default spread variables), we directly use the associated returns as a mimicking portfolio. Shanken (1992) has argued that this approach delivers the most efficient estimates of the risk premiums. When a factor is not itself an (excess) return (e.g., this is the case of macroeconomic variables such as industrial production growth, unexpected inflation, and real consumption growth), we construct the corresponding $K' \leq K$ mimicking portfolios by projecting the non-traded factors onto the space of excess returns of base assets and a set of control (predictive) variables ($j = 1, \ldots, K'$):

$$F_{j,t} = a_j + b_j' x_t + c_j' z_{t-1} + \varepsilon_{j,t} \qquad \varepsilon_{j,t} \sim N(0, \omega), \tag{13}$$

where $x_t$ is a vector of excess returns on the base assets (in this case, all defined to be zero investment portfolios) and $z_{t-1}$ denotes a vector of instruments that have the ability to predict returns. The resulting returns on the $i$th factor mimicking portfolio (FMP henceforth) are then defined as $FMP_{j,t} = \hat{a}_j + \hat{b}_j' x_t$ and collect the fitted component of a factor that is unpredictable on the basis of past information and that at the same time may be replicated by trading base assets using weights estimated by $\hat{b}_j$. Note that the coefficients $a_j$ and $b_j$ do not need to add up to one because the base assets are zero-investment portfolios (see Lamont 2001). The base assets include six equity zero net investment portfolios with different book-to-market and size characteristics as well as the returns on long-term government bonds minus the returns on the short term government bonds and the return on long-term corporate bonds minus the return on long-term government bonds. We choose these assets for their well known ability to span large "portions" of the return space. The set of instruments includes the lagged yield spread of long-term Treasury bonds minus the T-bill yield, the lagged yield spread of long-term corporate bonds minus the yield on long-term government bonds, and the lagged real short-term bill rate.

### 3.2 Prior Choices and Sensitivity

Realistic values for the different prior distributions obviously depend on the problem at hand. In testing factor models, priors calibration is particularly important because help to identify the factor exposures, which is otherwise rather problematic because the well-known indeterminacy problems upon rotations of factors and risk premia (see e.g., McCulloch and Rossi 1991 and Geweke and Zhou 1996). Table 2 provides an overview of prior hyper-parameters calibrations. In general, we use weak priors, excluding the size of the breaks and the break probabilities for which our priors are quite informative. We set $a_{ij} = 3.2$, $b_{ij} = 60$ and $\gamma_{ij} = 0.5$, $\delta_{ij} = 100$ which suggests that we can expected a relatively low probability of having a break with a moderate expected size. As far as the (log of) idiosyncratic risk is concerned, we assume a low break probability with $a_{iv} = 1$, $b_{iv} = 99$, while the expected size of the break is $\gamma_{iv} = 0.2$, $\delta_{iv} = 50$. These small prior probabilities makes the modeling dynamics rather parsimonious. The prior beliefs on the size of the breaks are inverse-gamma distributed. The scale parameters $\gamma_{ij}, \gamma_{iv}$ and the $\delta_{ij}, \delta_{iv}$ degrees of freedom reflect prior belief that shocks to risk exposures are larger than stochastic breaks in the dynamics of idiosyncratic volatility.

Table 2: Prior Choices for the B-TVB-SV Model

| Parameter | Values | Parameter | Values |
|---|---|---|---|
| $a_{ij}$ | 3.2 for $j = 1, .., K$ , $i = 1, ..., N$ | $a_{iv}$ | 1 for $i = 1, ..., N$ |
| $b_{ij}$ | 60 for $j = 1, .., K$ , $i = 1, ..., N$ | $b_{iv}$ | 99 for $i = 1, ..., N$ |
| $\gamma_{ij}$ | 0.5 for $j = 1, .., K$ , $i = 1, ..., N$ | $\gamma_{iv}$ | 0.2 for $i = 1, ..., N$ |
| $\delta_{ij}$ | 100 for $j = 1, .., K$ , $i = 1, ..., N$ | $\delta_{iv}$ | 50 for $i = 1, ..., N$ |
| $\underline{\lambda}$ | 0 | $\psi_0$ | 0.1 |
| $\underline{V}$ | 1e3 | $\Psi_0$ | 10 |

The posterior results for our B-TVB-SV model are not very sensitive to the prior settings for the hyper-parameters that govern the prior break probabilities, as is illustrated in a simple sensitivity analysis in Appendix C on both a simulated and the original dataset of the paper. However, Appendix C also shows that a different prior setting for the size of breaks could potentially impact the different posterior estimates substantially. As a whole, the results suggest that posterior estimates of break probabilities are hardly affected by the corresponding prior setting. Furthermore, the prior setting for the expected size of the breaks influence the posterior results of the timing and size of the breaks. Proper parametrization of the break size prior distributions are indeed of crucial importance for our empirical application. In order to mitigate the impact of prior belief on posterior estimates we used the initial ten years of the sample (i.e. 1972:01-1982:01) to empirically elicit the prior distributions.

Posterior densities for each model are based on 10,000 simulations of the Markov Chain Monte Carlo (MCMC) sampler outlined in Section 2. Of these, we take as a burn-in sample 2,000 draws storing every other of them to simulate the posterior distribution of parameters and latent variables. The resulting autocorrelations of the draws are very low. A convergence analysis in Appendix B shows that this guarantees accurate inference in our factor model.

# 4 An Empirical Application to the U.S. Cross-Section of Financial Returns

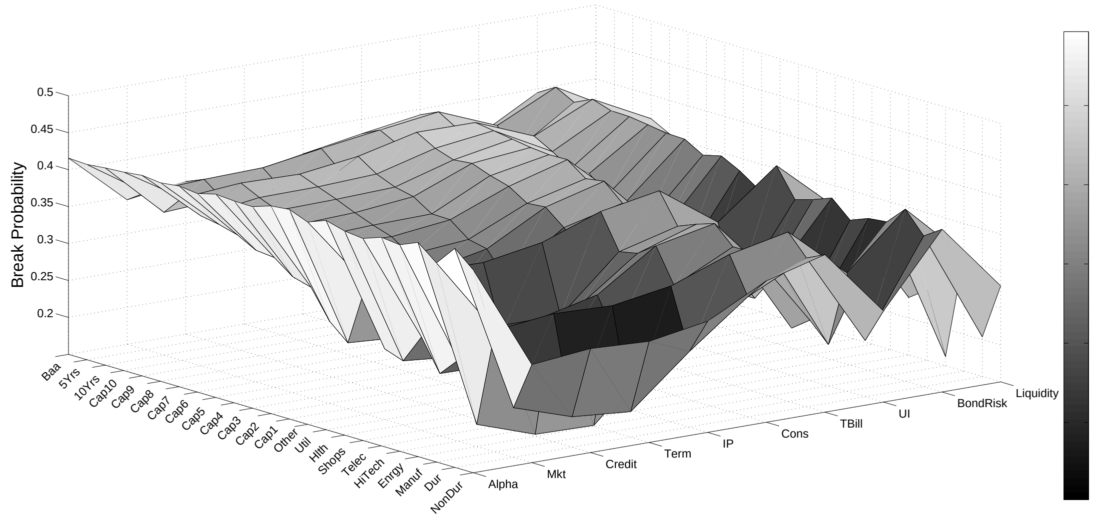
In this section, we focus on our empirical application: testing a macro-based factor pricing model both in the time series and in the cross-section of U.S. financial returns. The performance of our B-TVB-SV model are compared to a benchmark Fama-MacBeth two-steps procedure, as well as against a set of alternative model restrictions outlined in Section 2.1.

## 4.1 Time-Varying Betas

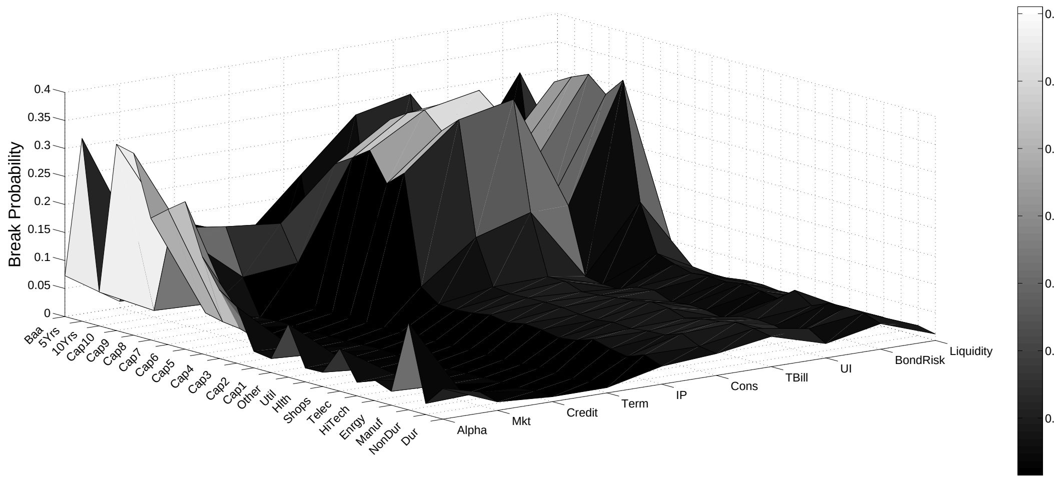
As an initial way to assess the plausibility of our results, Figure 1 reports the average (of posterior medians over time) probabilities over our sample of observing a break in the factor loadings, in addition to the intercept, across two different specifications, namely the B-TVB-SV and the homoskedastic B-TVB, for the 23 test assets/portfolios. Clearly the presence of breaks in the idiosyncratic variance process makes a difference in capturing any instability in portfolio betas. Under the B-TVP-SV model the average probability of observing a break is around 40% for the intercept (labeled as *alpha* in the figure) of all portfolios examined, and ranges from 20% for the credit and term spreads to almost 40% for the bond factor. This shows that infrequent and large breaks in betas (as well as Jensen's alphas) are often isolated by the Gibbs sampling algorithm. Under the B-TVB specification, instead, the degree of instability in the factor loadings dramatically collapses. The average probability of a break in betas is around 5% across all risk for the industry portfolios, while for both the size-sorted equity portfolios and bonds, the average break probability over the sample is between 20% and 30% across factors.

Figure 1: Mean Posterior Probability of Breaks in Factor Loadings Across Assets/Portfolios

This figure reports the posterior median estimates if breaks probabilities in betas across portfolios and factors for both the B-TVB-SV and B-TVB models. The sample period is 1972:01 - 2011:12. The first ten years of data are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The heating map is reported on the right-hand side.




(a) Bayesian Dynamic Model with Instability in Betas and Conditional Volatility



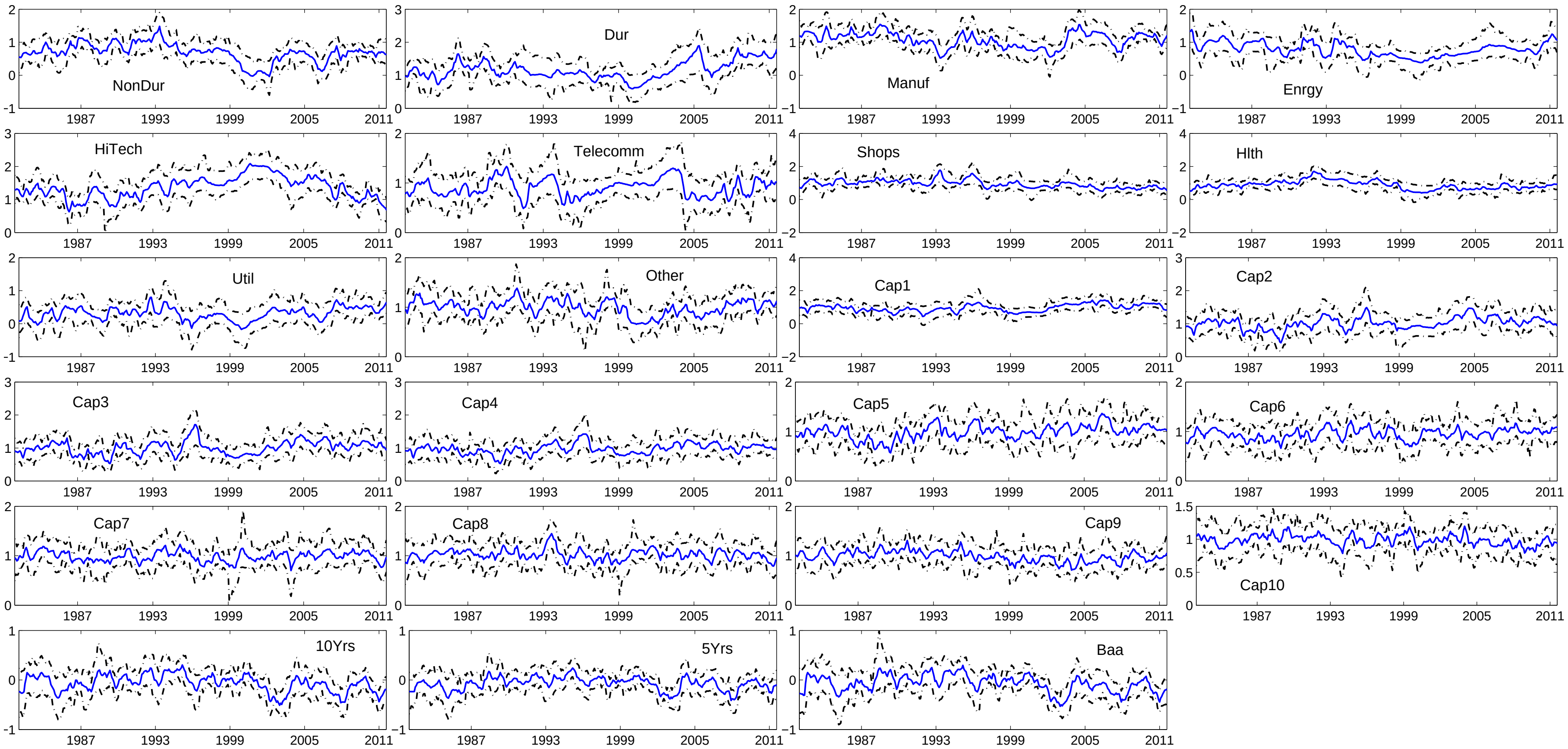

(b) Bayesian Dynamic model with Constant Conditional Volatility

Figures 2-6 plot a selection of time series medians and 95% Bayesian credibility intervals computed from the posterior densities of the loadings $\beta_{ij,t}$, obtained from the B-TVB-SV model. To save space, we report plots of time series of risk exposures for all the 23 portfolios used in our estimation, but only for three out of nine specific factors: the U.S. market portfolio, industrial production growth, and unexpected inflation. Other, similar plots concerning the remaining risk factors—the credit spread, the term spread, the real T-Bill, the real consumption growth, the bond and the liquidity factors —are available upon request even though we summarize their contents and implications below. An overview of the plots immediately reveals that the Bayesian estimates of the loadings for all but the market portfolio and the bond risk factor, imply a time path of the factor loadings that is rather smooth over time. This is a first interesting result: even though (6) formally allows factor exposures to be subject to "jumps" over time, as a result of the realization of $\kappa_{ij,t}$, the resulting posterior densities are actually smooth. Interestingly, this smoothness mimics exactly what many earlier papers have imposed by assuming either random walk or highly persistent stationary processes with small variance of the shocks, but is derived endogenously and is data-driven, which means that occasional large jumps in exposures and/or high volatility of the process may be accommodated. Second, with a limited number of exceptions that will noted below, the 95% confidence bands are relatively tight, which means that the betas are estimated with a fairly high level of reliability.

In particular, Figure 2, concerning exposures to market risk, collects most of the loadings for which we have evidence that betas are non-zero. All equity portfolios are characterized by positive and reliably estimated betas. This is not the case for the bond portfolios which essentially show zero exposure to the market risk factor. Figure 3 offers an opportunity to compare the B-TVB-SV estimates with market beta exposures under a the classical Fama-MacBeth approach. The plots of time-varying exposures to real output (industrial production growth) risk in Figure 4 show occasionally larg(er) 95% credibility regions that tend to widen over the sample. However, also in this case, for a large sub-set of portfolios, the corresponding betas are estimated to be negative and significant (nondurables, durables, manufacturing, high-tech, shops, health, and small- and medium-size equity portfolios), while for other portfolios the exposure is positive and significant (energy and utility stocks). Of course, negative exposures to output risk are partially surprising, but because in our model, factors have not been orthogonalized one vs. the others—that will require selecting and imposing a triangular structure

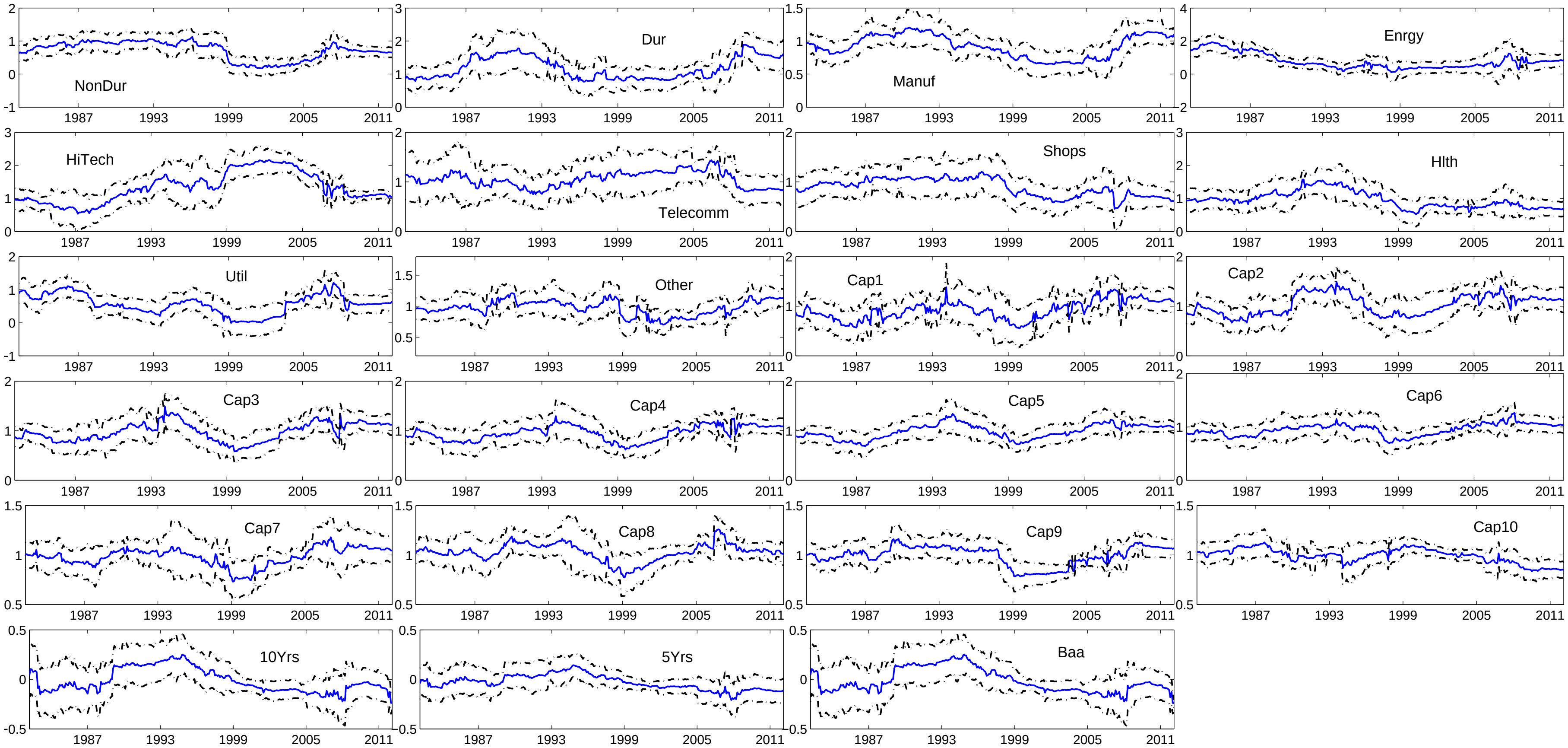
Figure 2: B-TVB-SV Factor Loadings: VW Market Portfolio

This figure reports the time series of the posterior median loadings for the market risk factor estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first ten years of data are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The solid blue line represents the posterior median. The dot-dashed black lines represents the 95% credibility interval.



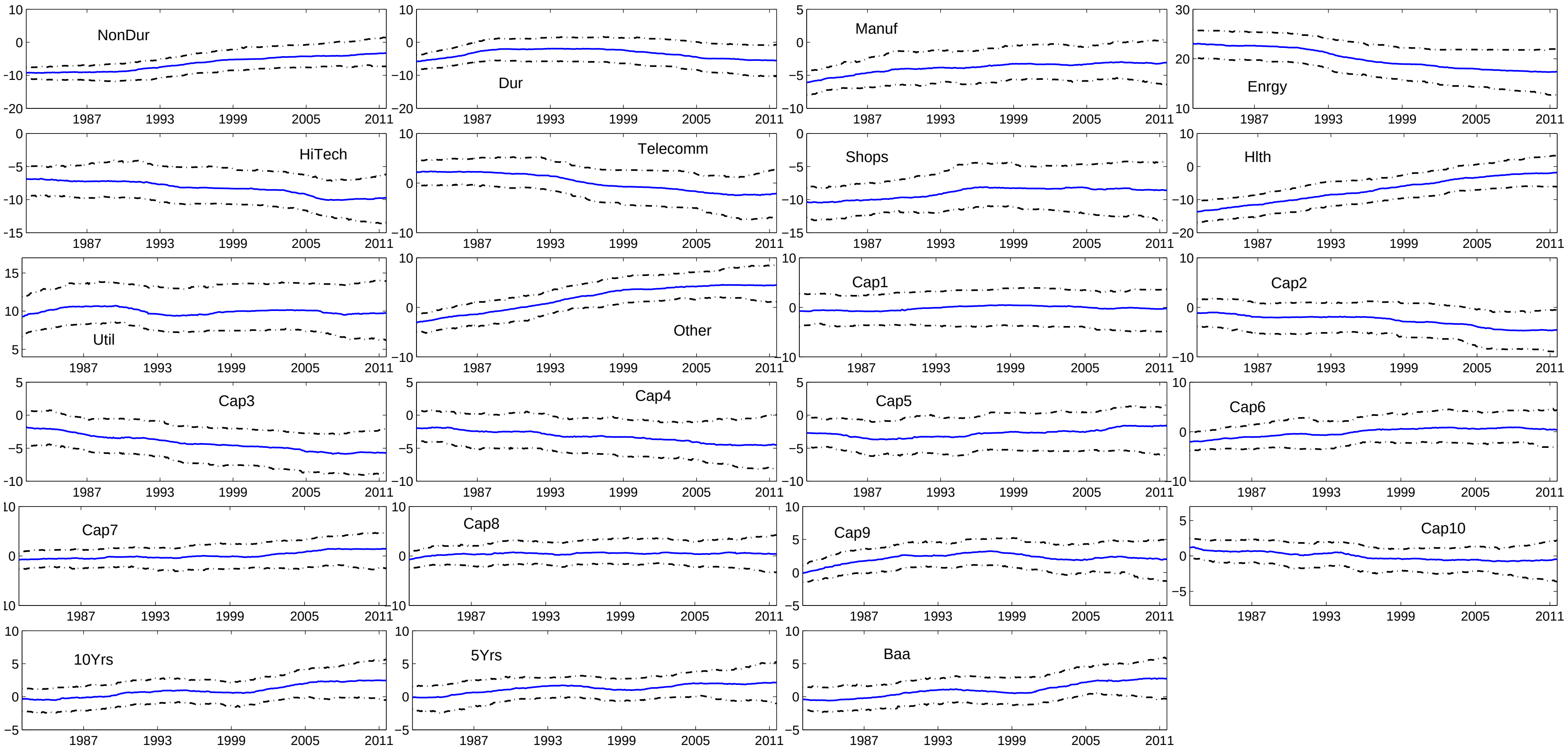
## Figure 3: Factor Loadings Estimated with the Fama-MacBeth Approach: VW Market Portfolio

This figure reports the time series of the of the posterior mean loadings for the market risk factor estimated from a 5-year rolling-window estimation approach. The sample period is 1972:01 - 2011:12. The solid blue line represents the posterior median. The dot-dashed black lines represents the 95% credibility interval. Asymptotic standard errors are computed assuming absence of cross-sectional dependence among the betas estimates.



# Figure 4: B-TVB-SV Factor Loadings: Industrial Production

This figure reports the time series of the posterior median loadings for the industrial production growth factor estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first ten years of data are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The solid blue line represents the posterior median. The dot-dashed black lines represents the 95% credibility interval.



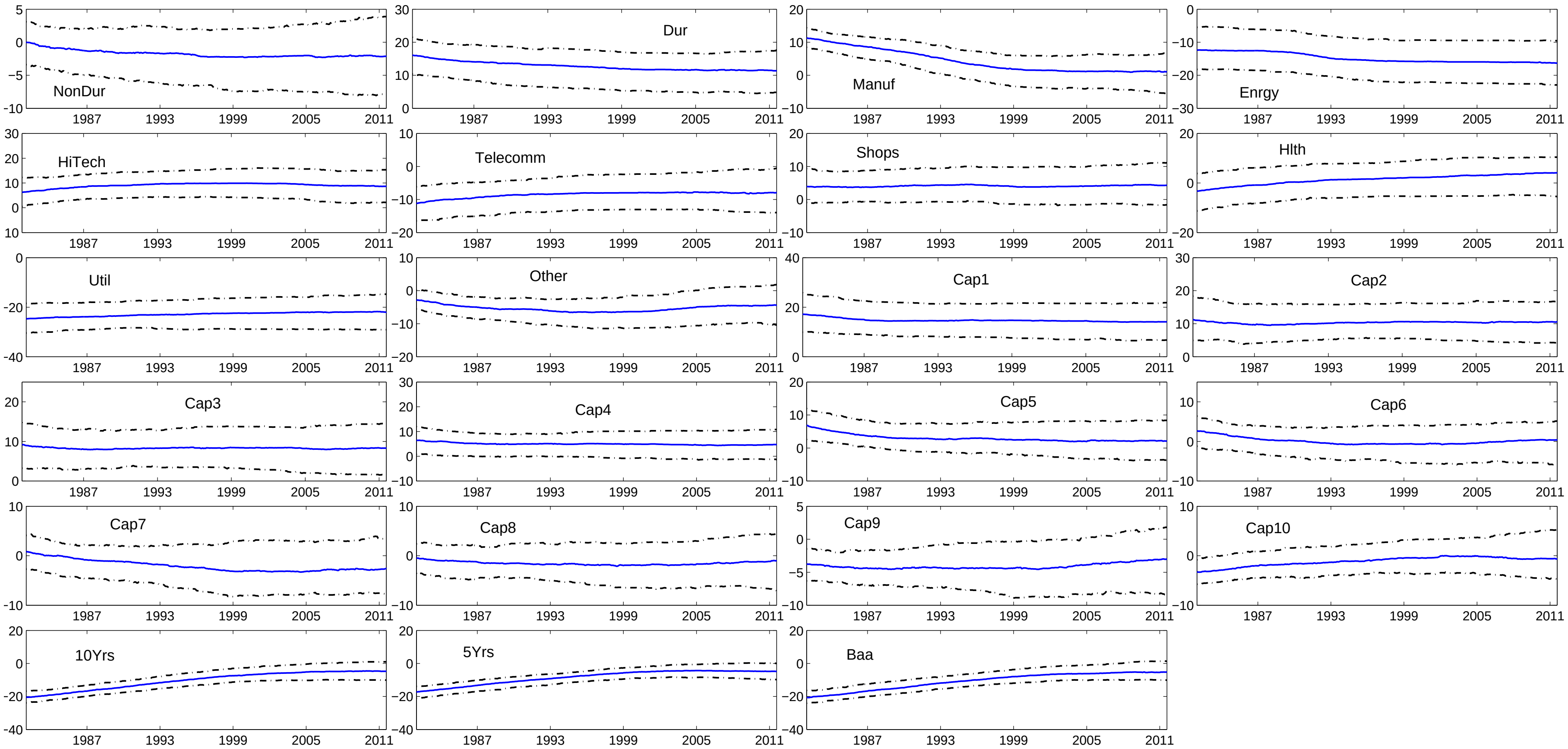
that would prove to be "ad-hoc" —betas only capture partial effects, after other exposures to business cycle risks are taken into account (see Kramer 1994). An unreported figure concerning betas vs. the short term real rate shows instead exposures that are small and for which the 95% credibility bands tend to include zero for most of the sample.

Figure 5 shows estimated time-varying exposures to unexpected inflation risks. In the asset pricing literature, the issue of the exposure of asset returns to inflation risks has often been debated. The plots show that even though confidence bands tend to be wider for this factor than for other factors that we have described before, for many portfolios there tends to be still significant evidence of a significantly positive exposure, i.e., of the fact that these assets pay out risk premia to compensate for inflation risks. Even if we limit ourselves to global results that hold throughout our entire sample, this hedging property obtains in the case of durables, high-tech, retail, and of small and medium-capitalization stocks. On the contrary, energy, telecommunication, utilities, and especially all kinds of bonds (including corporate junk), imply negative, significantly estimated exposures throughout the sample. An unreported figure concerning the exposures on default and real consumption growth risks, all betas imply low variability and narrow 95% credibility regions, but these also fluctuate steadily around zero for the *all* 23 portfolio investigated.

Figure 6 reports posterior medians and 95% credibility intervals for the $\beta_{i0,t}$s estimated from the B-TVB-SV model. In an ICAPM setting, when all factors are traded, $\beta_{i0,t}$ plays a key role: with $F_{j,t} = 0$ for $j = 1, ..., K$, then (1) simplifies to $r_{i,t} = \beta_{i0,t} + \epsilon_{i,t}$ (with $\epsilon_{i,t} \sim N(0, \sigma^2_{i,t})$) and any $E[\beta_{i0,t}] \neq 0$ would imply that in the absence of any priced risk factors, the excess return on asset/portfolio $i$ is not zero, which represents a violation of standard economic principles. Out of 23 portfolios, in no case the estimated mis-pricing indicators are systematically elevated (in absolute) value. In fact, apart from occasional fluctuations, separate calculations show that the 95% credibility regions include a zero mis-pricing in more than three-quarters of our sample. This is an indication that our B-TVB-SV the model is well-specified in an economic sense, as it does not imply any evidence of a systematic mis-pricing. Of course, in the case of many portfolios, occasional periods in which the posterior of $\beta_{i0,t}$ fails to include a zero mis-pricing can be found. For instance, there is evidence that all bond portfolios implied positive and tight posteriors for the Jensen's alphas between 2000 and 2004; high-tech and telecommunication stocks were all giving large and significant alphas during the early- to mid-1990s.

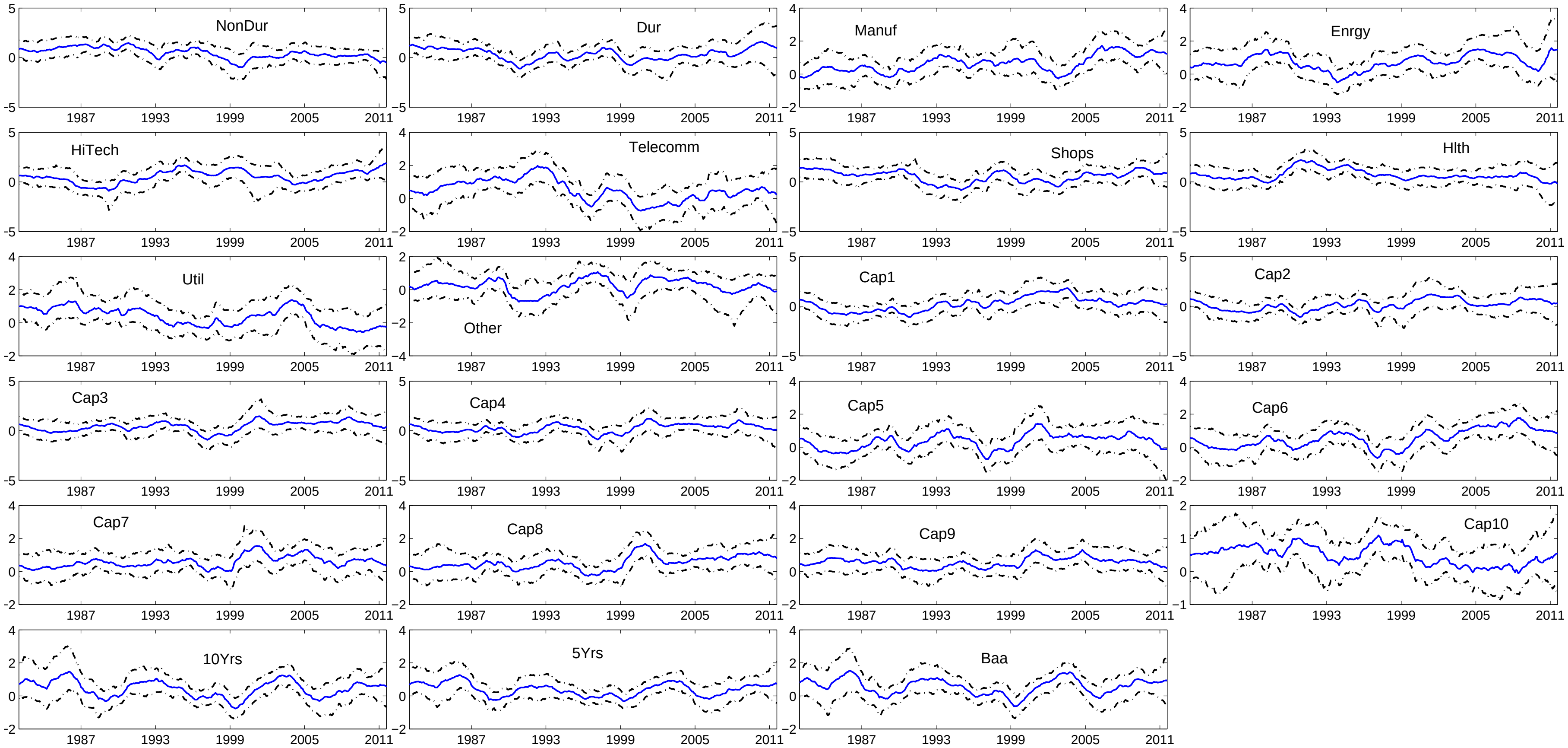
## Figure 5: B-TVB-SV Factor Loadings: Unexpected Inflation

This figure reports the time series of the posterior median loadings for the unexpected inflation factor estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first ten years of data are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The solid blue line represents the posterior median. The dot-dashed black lines represents the 95% credibility interval.



## Figure 6: B-TVB-SV Jensen's Alphas

This figure reports the time series of the posterior medians of the Jensen's alpha from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first ten years of data are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The solid blue line represents the posterior median. The dot-dashed black lines represents the 95% credibility interval.



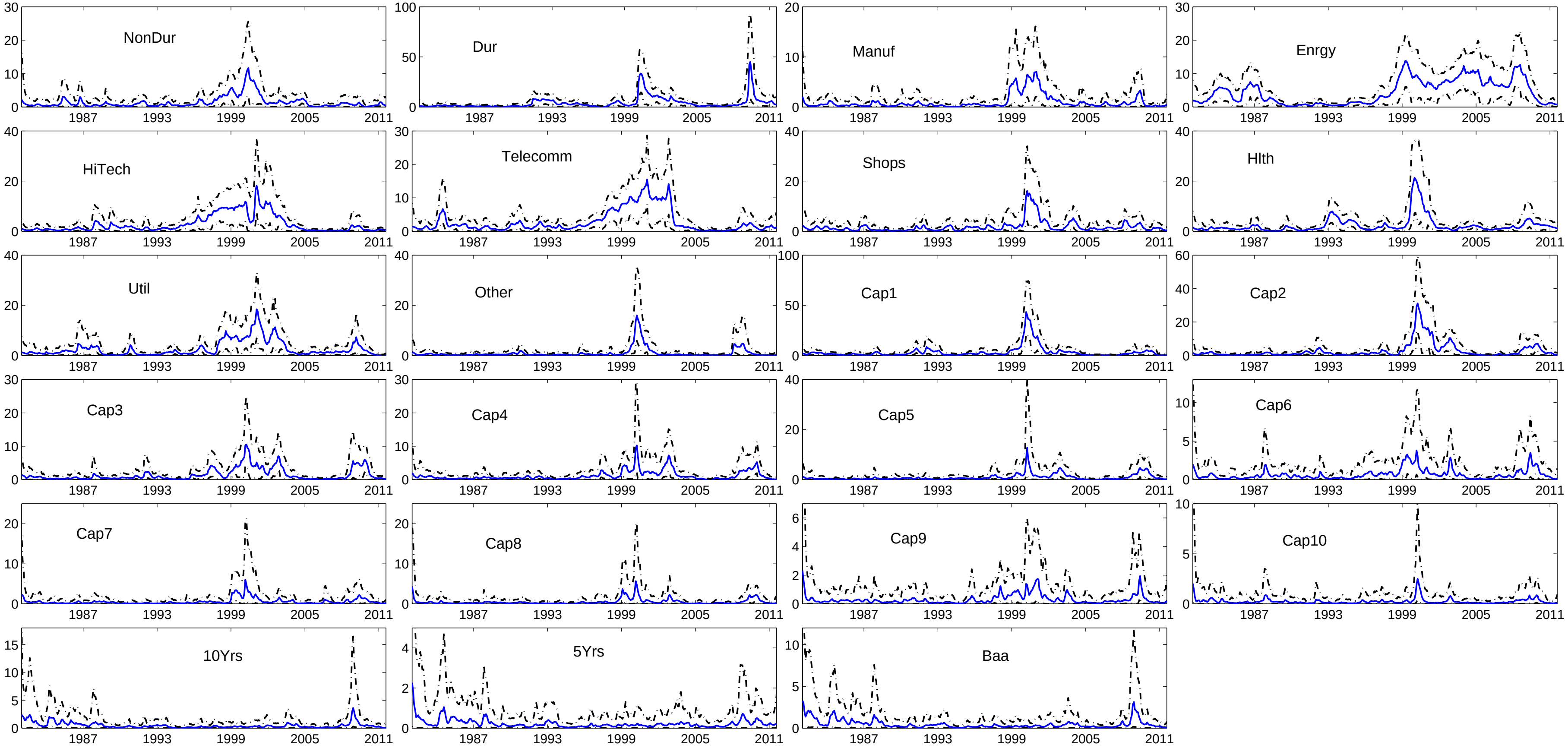
## 4.2 The Role of Idiosyncratic Risk

A growing literature (see e.g., Campbell et al. 2001) has stressed that the idiosyncratic variance of the excess returns on most test portfolios, $\sigma_{it}^2$, has undergone important shifts and/or dynamics over the last two decades. Figure 7 plots posterior medians for $\sigma_{it}^2$ estimated from our B-TVB-SV model, along with 95% credibility intervals. There are evident spikes in idiosyncratic volatility in the early 2000s and weaker signs of a growing trend towards the end of our sample. The financial crisis of 2008-2009 induces a residual risk increase, but this appears to be minor compared to 1999-2001, when the model had temporarily lost its ability to fit the U.S. cross-section. In this respect, the fact that the model is more at trouble with the tech stock bust than with the U.S. sub-prime and credit crunch crises is intriguing. However, the fact that idiosyncratic is counter-cyclical was largely expected in the light of the literature (see Campbell et al. 2001). The B-TVB-SV model explains away almost all the variability in excess returns in the case of medium and large cap stocks, and to some extent also government bonds. Spikes in idiosyncratic risk are instead more pronounced for small caps and for a number of industry portfolios, that are explained much less accurately than size-sorted portfolios are. Yet, no clear trend is observed, which is consistent with the more recent evidence reported by Bekaert et al. (2012).

Figure 1 shows that the presence of breaks in the idiosyncratic variance process makes a difference in capturing any instability in portfolio betas. The question now could be if the functional form imposed in (7) might indeed sensibly drive the results. A number of studies in the finance literature have compared alternative models of time-varying volatility of asset returns (e.g. Hansen and Lunde 2005, Geweke and Amisano 2010, and Clark and Ravazzolo 2014). More recently, Eisenstat and Strachan (2014) discuss estimation of volatility in the context of inflation, in particular whether it should be modeled as a stationary process or as a random walk. In our B-TVB-SV model when a break arrives, log-volatility follows a random walk. While such random walk assumption might be indeed useful for practical reasons, it can be criticized as inappropriate since it implies that the range of possible values of volatility is unbounded in probability in the limit, which is obviously something we do not observe in financial markets. On the other hand, stationary processes, say an AR(1), are bounded in the limit, even though are close to be non-stationary at monthly frequencies. In Appendix D we show that by fitting a highly persistent AR(1) instead of our change-point process does not

Figure 7: B-TVB-SV Idiosyncratic Risk Dynamics

This figure reports the time series of the posterior medians for idiosyncratic risk estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first ten years of data are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The solid blue line represents the posterior median. The dot-dashed black lines represents the 95% credibility interval.



sensibly change the dynamics of estimated idiosyncratic risks. In fact, a simple log-marginal likelihood model assessment shows that a parsimonious change-point process for $\sigma_{it}^2$ is preferred in finite samples as opposed to highly persistent, albeit stationary, AR(1) processes.

### 4.3 Risk Premia

Table 3 reports summary statistics for the sample estimates of the risk premia $\{\hat{\lambda}_{j,t}\}$ ($j = 1, ..., K$) from our B-TVB-SV model as opposed to the benchmark F-MB and the alternative model restrictions, namely the B-TVP and the B-TVB models. Top panel shows (frequentist) estimates from the second-pass F-MB approach. Clearly the classical estimation procedure that non-parametrically tracks time-variation in the parameters using 5-year rolling windows delivers economically weak implications: only two factors were accurately priced in the cross-section (the market and bond factors), but the former with a p-value exceeding the standard 0.05 threshold and the latter with a rather difficult, negative sign; moreover, the time series mean of $\hat{\lambda}_{0,t}$ turns out to be large, positive (0.29% per month), and statistically significant (its p-value is 0.034), which is problematic to our MFAPM because a non-zero average $\lambda_{0,t}$ implies that omitted risk factors with non-zero risk premia must be absorbed by the residual mean.

Therefore the background to our dynamic, state-space results is that a simple, *ad-hoc* rolling window implementation of the MFAPM in (1)-(2) would yield an embarrassing rejection of the model, in spite of the fact that we are employing as many as nine factors, some of them coming with a strong endorsement of cross-sectional explanatory power from the asset pricing literature. The second panel in Table 3 shows the estimates from the B-TVB-SV model. For the sake of comparison with the F-MB procedure, we report the sample properties of the *median* estimates of risk premia from the B-TVB-SV model. Note, this frequentist-like approach to test the cross-sectional implication of the model might indeed leave aside some important information on the posterior distribution of risk premia. However, the idea of Table 3 is to compare our B-TVB-SV framework and the F-MB approach within the same setting, while keeping at the background the fact that under the B-TVB-SV lambdas and betas are estimated jointly, fully acknowledging their unstable and uncertain nature. Table 3 gives evidence of precisely estimated market, liquidity, and macroeconomic (as capture by IP growth shocks) risk premia, with the correct, positive signs (0.339, 0.317, and 0.002 percent per month/unit of risk, respectively). Also the unexpected inflation risk premium is precisely estimated but with a negative sign, similarly to

Table 3: Risk Premia

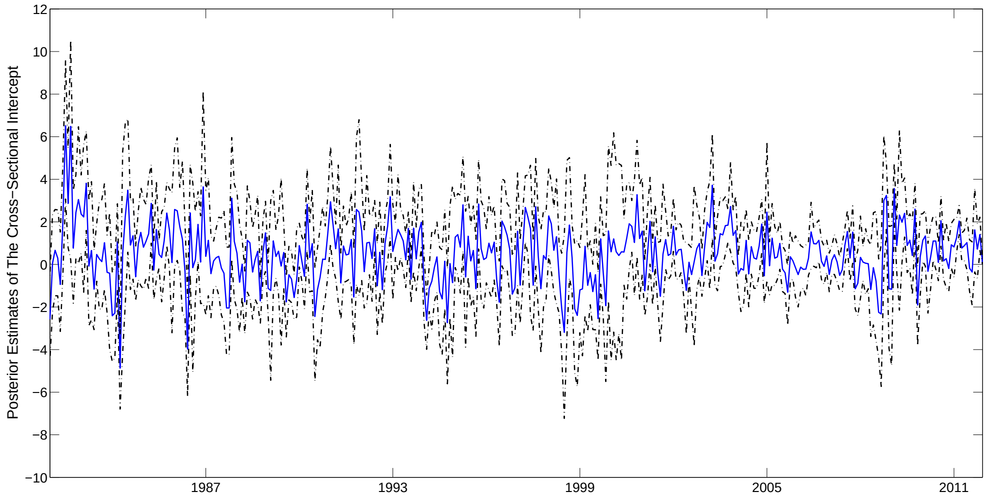
This table reports statistics describing the posterior distribution of the risk premia on each factor across different model specifications. *B-TVB-SV* stands for Bayesian time-varying betas, stochastic volatility model, while *B-TVB* and *B-TVP* are the dynamic Bayesian model restricted to have constant conditional volatility and random-walk betas, respectively. *F-MB* is the standard two-step procedure. Data are monthly and cover the sample period 1972:01 - 2011:12. The first ten years of monthly data are used to calibrate the priors for all the models except for the standard two-step Fama-MacBeth procedure.

| | **Average** | **Std. Error** | **t-stat** | **p-value** | **2.5%** | **50%** | **97.5%** |
|---|---|---|---|---|---|---|---|
| | | | **Two-step F-MB** | | | | |
| **Intercept** | 0.2909 | 0.1363 | 2.1350 | **0.0336** | -3.3346 | 0.3281 | 3.3471 |
| **Market** | 0.2593 | 0.1408 | 1.8414 | **0.0665** | -8.7553 | 0.6739 | 7.9319 |
| **Credit Spread** | 0.2208 | 0.2706 | 0.8161 | 0.4151 | -4.5672 | 0.3022 | 4.8480 |
| **Term spread** | 0.0042 | 0.0347 | 0.1201 | 0.9045 | -1.0198 | -0.0018 | 1.1440 |
| **IP Growth** | -0.0130 | 0.0092 | -1.4086 | 0.1600 | -0.3368 | -0.0210 | 0.3218 |
| **Real Consuption Growth** | 0.0061 | 0.0039 | 1.5485 | 0.1226 | -0.1917 | -0.0006 | 0.2309 |
| **Real T-bill Rate** | -0.0264 | 0.0412 | -0.6414 | 0.5217 | -1.4068 | 0.0145 | 1.4703 |
| **Unexpected Inflation** | -0.0085 | 0.0062 | -1.3670 | 0.1727 | -0.2079 | -0.0134 | 0.2100 |
| **Bond Risk Factor** | -0.4633 | 0.1883 | -2.4598 | **0.0145** | -6.9685 | -0.3831 | 5.1747 |
| **Liquidity Factor** | 0.4012 | 0.3471 | 1.1558 | 0.2487 | -12.4488 | 0.1301 | 12.9321 |
| | | | **B-TVB-SV** | | | | |
| **Intercept** | 0.4125 | 0.2924 | 1.4108 | 0.1593 | 0.3406 | 0.4913 | 0.6432 |
| **Market** | 0.3391 | 0.1298 | 2.6119 | **0.0095** | 0.1207 | 0.3482 | 0.5515 |
| **Credit Spread** | -0.1339 | 0.1145 | -1.1688 | 0.2434 | -0.0471 | 0.1291 | 0.3172 |
| **Term Spread** | -0.0149 | 0.0306 | -0.4880 | 0.6259 | -0.0334 | 0.0144 | 0.0616 |
| **IP Growth** | 0.0190 | 0.0076 | 2.4940 | **0.0132** | 0.0031 | 0.0188 | 0.0231 |
| **Real Consuption Growth** | 0.0020 | 0.0044 | 0.4575 | 0.6476 | -0.0054 | 0.0018 | 0.0090 |
| **Real T-bill Rate** | 0.0199 | 0.0300 | 0.6616 | 0.5087 | -0.0279 | 0.0187 | 0.0682 |
| **Unexpected Inflation** | -0.0206 | 0.0064 | -3.2095 | **0.0015** | -0.0211 | -0.0148 | -0.0007 |
| **Bond Risk Factor** | -0.0259 | 0.0719 | -0.3605 | 0.7187 | -0.1449 | -0.0218 | 0.0916 |
| **Liquidity Factor** | 0.3172 | 0.1560 | 2.0341 | **0.0428** | 0.0312 | 0.3214 | 0.5719 |
| | | | **B-TVP** | | | | |
| **Intercept** | 0.5862 | 0.0787 | 7.4482 | **0.0000** | 0.4575 | 0.5889 | 0.7172 |
| **Market** | 0.2197 | 0.0988 | 2.2237 | **0.0269** | 0.0472 | 0.2220 | 0.3786 |
| **Credit Spread** | 0.0139 | 0.0919 | 0.1516 | 0.8796 | -0.1381 | 0.0100 | 0.1710 |
| **Term Spread** | 0.0030 | 0.0213 | 0.1401 | 0.8887 | -0.0312 | 0.0020 | 0.0382 |
| **IP Growth** | -0.0079 | 0.0075 | -1.0473 | 0.2958 | -0.0209 | -0.0077 | 0.0036 |
| **Real Consuption Growth** | 0.0047 | 0.0040 | 1.1780 | 0.2397 | -0.0013 | 0.0046 | 0.0114 |
| **Real T-bill Rate** | 0.0067 | 0.0216 | 0.3094 | 0.7573 | -0.0278 | 0.0055 | 0.0424 |
| **Unexpected Inflation** | -0.0092 | 0.0054 | -1.6933 | **0.0914** | -0.0183 | -0.0090 | -0.0001 |
| **Bond Risk Factor** | -0.0126 | 0.0492 | -0.2550 | 0.7989 | -0.0982 | -0.0093 | 0.0666 |
| **Liquidity Factor** | 0.2071 | 0.1050 | 1.9726 | **0.0495** | 0.0325 | 0.2066 | 0.3720 |
| | | | **B-TVB** | | | | |
| **Intercept** | 0.5550 | 0.2775 | 1.9996 | **0.0464** | 0.2500 | 0.5421 | 0.8540 |
| **Market** | 0.3006 | 0.1448 | 2.0758 | **0.0388** | 0.0291 | 0.3028 | 0.5814 |
| **Credit Spread** | 0.1162 | 0.1582 | 0.7346 | 0.4631 | -0.1080 | 0.0967 | 0.3963 |
| **Term Spread** | 0.0130 | 0.0550 | 0.2368 | 0.8130 | -0.0812 | 0.0105 | 0.1120 |
| **IP Growth** | -0.0067 | 0.0101 | -0.6573 | 0.5115 | -0.0218 | -0.0060 | 0.0111 |
| **Real Consuption Growth** | 0.0030 | 0.0063 | 0.4788 | 0.6324 | -0.0072 | 0.0026 | 0.0141 |
| **Real T-bill Rate** | 0.0191 | 0.0498 | 0.3837 | 0.7014 | -0.0620 | 0.0165 | 0.1016 |
| **Unexpected Inflation** | -0.0024 | 0.0080 | -0.3066 | 0.7594 | -0.0172 | -0.0021 | 0.0103 |
| **Bond Risk Factor** | 0.0431 | 0.1119 | 0.3847 | 0.7007 | -0.1242 | 0.0313 | 0.2512 |
| **Liquidity Factor** | 0.0474 | 0.2419 | 0.1958 | 0.8449 | -0.3394 | 0.0229 | 0.4512 |

Chen et al. (1986), Ferson and Harvey (1991), and Lamont (2001). Importantly, the average posterior median across the sample for $\lambda_{0,t}$ is significantly different from zero. Figure 8 shows the whole posterior distribution of $\lambda_{0,t}$ across the sample. Except for occasional, short-lived nuances, there is no systematic evidence of a statistically significant $\lambda_{0,t}$ in our sample.

Figure 8: B-TVB-SV Cross-Sectional Intercept

This figure reports the time series of the posterior medians for the cross-sectional intercept $\lambda_{0,t}$ estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first ten years of data are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The solid blue line represents the posterior median. The dot-dashed black lines represents the 95% credibility interval.




All in all, these results illustrate the fact that while in a naive F-MB implementation all one gets is evidence that a standard multi-factor model—both in terms of its structure and for what concerns the factor it includes—is rejected with reference to a wide but typical set of U.S. financial asset portfolios, such finding is replaced by reassuring evidence that not only the market portfolio (as typical of textbook CAPM) but also a number of macroeconomic factors carry precisely estimated and economically meaningful risk prices. Such empirical findings are less comforting when we impose restrictions on the B-TVP-SV model. In both cases, the average posterior median of $\lambda_{0,t}$ is significantly positive, with p-values below 0.05. While in the B-TVP case at least 3 of the 4 factors that commanded positive and significant risk premia in B-TVP-SV set up, in the homoskedastic B-TVB model only market risks appear to be barely priced in the U.S. cross-section of stocks and bonds. This is indicative of the restrictions imposed by the B-TVP and the homoskedastic B-TVB models being rejected, an aspect that shall be investigated in more detail in Section 4.4.

### 4.4 Discriminating Among Models: Marginal Likelihood Evidence

Following McCulloch and Rossi (1991), we use the marginal likelihood of different models to perform a comparison able to take into account their overall (in-sample) statistical performance, and not only their asset pricing plausibility as in Sections 4.2-4.4. The marginal likelihood of a model is known to take into account both the uncertainty about the size and the presence of structural breaks and the uncertainty concerning the parameters in (4)-(8). The marginal likelihood of each model is computed as

$$p(R|F;\mathcal{M}_i) = \int ... \int \sum_{\mathcal{K}} p(R|B,\mathcal{K},\Sigma,\lambda,\theta,F;\mathcal{M}_i) \times p(\theta,B,\mathcal{K},\Sigma,\lambda|R,F;\mathcal{M}_i)dBd\Sigma d\theta, \quad (14)$$

where $\mathcal{M}_i$ identifies the $i$th model and $p(\theta,B,\mathcal{K},\Sigma,\lambda|R,F;\mathcal{M}_i)$ the posterior density from (12). Following Chib (1995), we compute the (log of the) marginal likelihood by replacing the unobservable breaks and parameters in the likelihood of the data generating process defined by (4) for each draw.

Table 4 reports the log-marginal likelihoods for each of the model specifications as well as the Bayes factors, the difference between model-specific (log) likelihoods, used as a model selection indicator that naturally penalizes for the different size/complexity of different models (see Kass and Raftery 1995), for each of the alternative frameworks including the two-step F-MB approach, vs. B-TVB-SV. Bayes factors measure a model ability to explain the entire distribution (not just first moments) of test asset returns, and therefore permit the simultaneous comparison of multiple models, regardless of whether the models are nested. To favor interpretations, we report the log-marginal likelihood contributions by each of the 23 test portfolios under each of the models. Interestingly, the B-TVB-SV model shows the higher log-marginal likelihood values across all of the portfolios under consideration. By exceeding 100, all the overall Bayes factors are highly significant. In particular, the factors vs. the B-TVP and the two-step F-MB implementations are on average 892 and 5602, respectively, and therefore appear to be decisively in favor of the complete B-TVB-SV framework. The Bayes factor vs. the B-TVB

Table 4: Marginal Likelihoods and Bayes Factors Across Alternative Model Specifications

This table reports the values of the log-marginal likelihoods and the relative Bayes Factors for different model specifications. The values reported are disaggregated by computing the contributions coming from each of the portfolios under investigation. *B-TVB-SV* stands for Bayesian time-varying betas, stochastic volatility model, while *B-TVB* and *B-TVP* are the dynamic Bayesian model restricted to have constant conditional volatility and random-walk betas, respectively. *F-MB* is the standard two-step procedure. *BF1* is the Bayes Factor for the B-TVB-SV model vs. the no-stochastic volatility restriction. Likewise, *BF2* and *BF3* are the Bayes Factors comparing the B-TVB-SV model with the B-TVP and the F-MB approaches.

| | **B-TVB-SV** | **B-TVB** | **B-TVP** | **F-MB** | **BF1** | **BF2** | **BF3** |
|---|---|---|---|---|---|---|---|
| | **10 Industry Portfolios, Value-Weighted** | | | | | | |
| **Non Durable Goods** | **-445.39** | -1408.71 | -635.40 | -3131.83 | 963.32 | 190.00 | 2686.44 |
| **Durable Goods** | **-700.77** | -1980.33 | -832.07 | -4412.78 | 1279.56 | 131.29 | 3712.01 |
| **Manufacturing** | **-330.98** | -1199.96 | -522.95 | -3851.11 | 868.98 | 191.97 | 3520.13 |
| **Energy** | **-789.61** | -1793.14 | -821.64 | -2687.83 | 1003.53 | 32.03 | 1898.22 |
| **High Tech** | **-571.53** | -1732.31 | -717.08 | -7269.27 | 1160.78 | 145.55 | 6697.74 |
| **Telecommunications** | **-614.18** | -1634.38 | -734.46 | -3353.11 | 1020.20 | 120.28 | 2738.93 |
| **Shops and Retail** | **-481.33** | -1370.61 | -648.98 | -4271.70 | 889.29 | 167.65 | 3790.38 |
| **Health** | **-613.64** | -1591.56 | -706.14 | -3107.84 | 977.92 | 92.50 | 2494.21 |
| **Utilities** | **-572.88** | -1684.43 | -698.85 | -1955.89 | 1111.55 | 125.96 | 1383.01 |
| **Other** | **-270.90** | -1345.87 | -519.80 | -6041.50 | 1074.97 | 248.90 | 5770.60 |
| | **10 Size-Sorted Portfolios, Value-Weighted** | | | | | | |
| **Decile 1** | **-620.66** | -1756.62 | -725.44 | -7211.50 | 1135.96 | 104.77 | 6590.84 |
| **Decile 2** | **-535.44** | -1632.89 | -678.94 | -6578.42 | 1097.46 | 143.50 | 6042.98 |
| **Decile 3** | **-428.11** | -1337.08 | -616.02 | -6506.86 | 908.96 | 187.90 | 6078.74 |
| **Decile 4** | **-392.01** | -1262.43 | -589.11 | -7127.84 | 870.43 | 197.11 | 6735.84 |
| **Decile 5** | **-335.93** | -1134.01 | -543.20 | -7517.14 | 798.08 | 207.27 | 7181.21 |
| **Decile 6** | **-259.32** | -932.83 | -506.59 | -8008.70 | 673.51 | 247.27 | 7749.38 |
| **Decile 7** | **-202.41** | -923.42 | -468.69 | -7121.70 | 721.01 | 266.28 | 6919.29 |
| **Decile 8** | **-149.14** | -882.35 | -446.04 | -8924.98 | 733.21 | 296.90 | 8775.84 |
| **Decile 9** | **-95.261** | -601.39 | -365.63 | -8158.70 | 506.13 | 270.37 | 8063.44 |
| **Decile 10** | **-54.063** | -544.78 | -329.60 | -7820.37 | 490.72 | 275.55 | 7766.31 |
| | **Bond Returns** | | | | | | |
| **10 - Yrs Treasury** | **-188.98** | -1052.55 | -422.69 | -9201.70 | 863.57 | 233.70 | 9012.72 |
| **5 - Yrs Treasury** | **-41.972** | -549.00 | -320.55 | -7951.90 | 507.03 | 278.58 | 7909.93 |
| **Baa Corporate Bonds (10-20 years)** | **-185.96** | -1050.90 | -420.63 | -5522.90 | 864.94 | 234.67 | 5336.94 |

model with stochastic volatility is instead 190 on average and remains favorable to B-TVB-SV. Surprisingly, the B-TVP model ranks second both in overall terms and for all the test portfolios, thus outperforming the homoskedastic B-TVB alternative. This result emphasizes that by fully acknowledging instability in idiosyncratic risk plays a key role beyond that of capturing breaks in the betas. As one would expect, given its ingenious but ad-hoc nature, the classical two-step F-MB approach ranks last with an overall marginal likelihood around 15 times lower than under the B-TVB-SV model. The dominance of the B-TVB-SV framework occurs across all portfolios, but appears to be particularly elevated in the case of bonds and medium and large caps portfolios of stocks.

# 5 Economic Assessment

So far our discussion has focused on the statistical performance in terms of whether there was evidence of either the $\lambda_{0,t}$s or the $\beta_{i0,t}$s coefficients being different from zero and with emphasis on the comparison of log-marginal likelihood values. We have concluded that (1)-(2) is rejected in its two-pass F-MB implementation based on 5-year rolling window estimates. However, there was some supportive indications that the B-TVB-SV model may be not completely at odds with the data. The results concerning B-TVP have shown that while there are some degrees of freedom as to the way one ought to best model time-variation in risk exposures, capturing instability in stochastic volatility is truly fundamental. Yet, we still know little about the economic implications of B-TVB-SV. In this section, we report additional evidence on the economic importance of the estimates uncovered for B-TVB-SV model.

## 5.1 Variance Ratios and Sources of Risks

With reference to the estimates of (4)-(8), we have computed (posterior distributions of the) variance ratios, $VR1$ and $VR2$ described in Appendix E. Such ratios measure the degree of mis-specification of a MFAPM. The idea of $VR1$ and $VR2$ is that a correctly specified MFAPM should at least explain most or all of the predictable variation in the excess returns of the test assets, and therefore leave an unexplained portion that should be as small as possible. Given their popularity, we just limit ourselves to recall that $VR1$ should be equal to 1 if the multi-factor model is correctly specified, which means that all the predictable variation in excess returns is captured by variation in macroeconomic risk; at the same time, $VR2$ should be equal to zero if the multi-factor model is correctly specified. Note, $VR1 = 1$ does not imply that $VR2 = 0$ and viceversa (see Appendix E). In what follows, the information at time $t-1$ ($\mathbf{Z}_{t-1}$) used to tease out the total predictable variation in excess returns used as a "denominator" in the variance ratios is proxied by the instrumental variables listed in Table 1, plus a dummy variable to account for the so-called "January effect" (see Thaler 1987).

Columns 4 and 7 of Table 5 present posterior medians of $VR1$ and $VR2$ obtained from the B-TVB-SV model for each of the 23 portfolios. These variance ratios are compared to the ones obtained from competing models. Variance ratio results are encouraging. Under a $VR1$ perspective, on average approximately 80% of the predictable variation in excess returns is captured by the B-TVB-SV model. Such a statistic is only 51% in the case of the F-MB implementation (column 1) and goes as low as 47 and 43% for the B-TVP and homoskedastic B-TVB models, respectively. Although in the light of the earlier log-marginal likelihood evidence, this is relatively un-surprising, these results remain economically meaningful. However, the generally high $VR1$ ratios from the B-TVB-SV model vary considerably across different test assets. The ratios are relatively high, also in relation to what is typically reported in the literature (see Ferson and Harvey 1991), in the case of government bond portfolios (possibly because we have used Cochrane and Piazzesi's factor) and for a few industries, such as manufacturing, energy, and high-tech, for which VR1 exceeds 90%. It is instead below 50% in the case of the smallest capitalization decile and of non-investment grade corporate bonds, exactly where one would expect our macroeconomic risk factors to have more trouble at fitting the variation in excess returns.

Because $VR1 + VR2 = 1$ does not hold (see Appendix E), the finding of high $VR1$ ratios fails to imply that the $VR2$ ratios are close to zero. Yet, $VR2$ is on average just above 20% in the case B-TVB-SV, to be contrasted with averages across test portfolios of 48-54% in the case of other models. Moreover, in the case of the B-TVB-SV framework, we record $VR2$ ratios equal to or inferior to 15% in 9 out of 23 portfolios. All in all, under both the $VR1$ and $VR2$, we find evidence of appreciable performance of the model.

Appendix E reports also further results on the contribution of each risk factor to fitting the predictable variation in excess stock returns. The highest contribution is given by the market risk factor: with three exceptions (energy, health, and utility stocks), all the ratios concerning

# Table 5: Variance Decomposition Tests Across Models

This table reports the results of variance decomposition tests across models. The first two columns show the values from the standard two-step Fama-MacBeth methodology. All rates are in excess of the holding period return on a 1-month T-Bill. VR1 is the ratio of the variance of a model predicted returns and the variance of expected returns estimated from a projection on a set of instruments $Z_t$. VR2 is the ratio of the variance of the predictable part of returns not explained by a model and the variance of projected returns. The instrumental variables are the lagged monthly dividend yield on the NYSE/AMEX, the lagged yield of a Baa corporate bond, and the lagged spread of long- vs. short-term government bond yields. *B-TVB-SV* stands for Bayesian time-varying betas, stochastic volatility model, while *B-TVB* and *B-TVP* are the dynamic Bayesian model restricted to have constant conditional volatility and random-walk betas, respectively. *F-MB* is the standard two-step procedure.

| | F-MB | | B-TVB-SV | | | | | | B-TVP | | | | | | B-TVB | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | VR1 | VR2 | VR1 | | | VR2 | | | VR1 | | | VR2 | | | VR1 | | | VR2 | | |
| | | | 2.5% | 50% | 97.5% | 2.5% | 50% | 97.5% | 2.5% | 50% | 97.5% | 2.5% | 50% | 97.5% | 2.5% | 50% | 97.5% | 2.5% | 50% | 97.5% |
| | | | | | | | | | **10 Industry Portfolios, Value-Weighted** | | | | | | | | | | | |
| Non Durable Goods | 0.368 | 0.695 | 0.318 | 0.454 | 0.530 | 0.089 | 0.144 | 0.294 | 0.254 | 0.320 | 0.450 | 0.241 | 0.328 | 0.418 | 0.046 | 0.217 | 0.749 | 0.243 | 0.719 | 0.835 |
| Durable Goods | 0.499 | 0.450 | 0.547 | 0.833 | 0.969 | 0.098 | 0.131 | 0.213 | 0.256 | 0.376 | 0.789 | 0.239 | 0.678 | 0.783 | 0.072 | 0.464 | 0.802 | 0.187 | 0.519 | 0.832 |
| Manufacturing | 0.717 | 0.205 | 0.756 | 0.916 | 1.053 | 0.031 | 0.173 | 0.309 | 0.090 | 0.453 | 0.816 | 0.142 | 0.418 | 0.934 | 0.134 | 0.492 | 0.818 | 0.200 | 0.569 | 0.845 |
| Energy | 0.698 | 0.354 | 0.897 | 0.987 | 1.090 | 0.017 | 0.056 | 0.067 | 0.707 | 0.832 | 0.952 | 0.115 | 0.221 | 0.305 | 0.613 | 0.865 | 0.938 | 0.087 | 0.148 | 0.296 |
| High Tech | 0.660 | 0.380 | 0.869 | 0.941 | 1.011 | 0.108 | 0.138 | 0.167 | 0.345 | 0.546 | 0.812 | 0.256 | 0.431 | 0.680 | 0.337 | 0.614 | 0.854 | 0.105 | 0.309 | 0.638 |
| Telecommunications | 0.494 | 0.476 | 0.730 | 0.888 | 0.978 | 0.109 | 0.154 | 0.212 | 0.117 | 0.436 | 0.918 | 0.073 | 0.636 | 0.828 | 0.176 | 0.418 | 0.706 | 0.326 | 0.533 | 0.799 |
| Shops and Retail | 0.663 | 0.336 | 0.615 | 0.795 | 0.913 | 0.045 | 0.213 | 0.378 | 0.112 | 0.383 | 0.901 | 0.139 | 0.529 | 0.922 | 0.190 | 0.410 | 0.832 | 0.159 | 0.561 | 0.790 |
| Health | 0.428 | 0.539 | 0.612 | 0.830 | 0.886 | 0.003 | 0.116 | 0.227 | 0.202 | 0.479 | 0.782 | 0.247 | 0.518 | 0.675 | 0.134 | 0.490 | 0.796 | 0.144 | 0.560 | 0.804 |
| Utilities | 0.266 | 0.705 | 0.432 | 0.600 | 0.690 | 0.017 | 0.410 | 0.760 | 0.066 | 0.320 | 0.869 | 0.134 | 0.520 | 0.910 | 0.101 | 0.402 | 0.708 | 0.266 | 0.592 | 0.871 |
| Other | 0.278 | 0.700 | 0.431 | 0.615 | 0.699 | 0.152 | 0.357 | 0.744 | 0.231 | 0.375 | 0.552 | 0.428 | 0.592 | 0.775 | 0.240 | 0.385 | 0.633 | 0.311 | 0.618 | 0.660 |
| | | | | | | | | | **10 Size-Sorted Portfolios, Value-Weighted** | | | | | | | | | | | |
| Decile 1 | 0.314 | 0.677 | 0.225 | 0.309 | 0.364 | 0.281 | 0.648 | 0.797 | 0.242 | 0.342 | 0.450 | 0.552 | 0.681 | 0.704 | 0.015 | 0.296 | 0.622 | 0.305 | 0.750 | 0.901 |
| Decile 2 | 0.731 | 0.182 | 0.723 | 0.882 | 0.985 | 0.067 | 0.159 | 0.255 | 0.141 | 0.467 | 0.842 | 0.190 | 0.564 | 0.761 | 0.278 | 0.456 | 0.689 | 0.218 | 0.473 | 0.704 |
| Decile 3 | 0.629 | 0.323 | 0.611 | 0.906 | 0.954 | 0.002 | 0.163 | 0.316 | 0.278 | 0.569 | 0.976 | 0.131 | 0.409 | 0.889 | 0.168 | 0.442 | 0.706 | 0.187 | 0.389 | 0.783 |
| Decile 4 | 0.603 | 0.380 | 0.542 | 0.820 | 0.961 | 0.043 | 0.178 | 0.408 | 0.270 | 0.520 | 0.886 | 0.234 | 0.516 | 0.814 | 0.239 | 0.488 | 0.906 | 0.079 | 0.476 | 0.847 |
| Decile 5 | 0.538 | 0.497 | 0.633 | 0.877 | 0.999 | 0.060 | 0.112 | 0.311 | 0.089 | 0.465 | 0.951 | 0.080 | 0.473 | 0.820 | 0.061 | 0.261 | 0.552 | 0.471 | 0.750 | 0.906 |
| Decile 6 | 0.262 | 0.679 | 0.620 | 0.864 | 1.012 | 0.030 | 0.166 | 0.348 | 0.070 | 0.418 | 0.824 | 0.120 | 0.500 | 0.899 | 0.154 | 0.354 | 0.737 | 0.198 | 0.635 | 0.891 |
| Decile 7 | 0.448 | 0.515 | 0.723 | 0.846 | 0.993 | 0.013 | 0.183 | 0.270 | 0.102 | 0.478 | 0.858 | 0.144 | 0.360 | 0.813 | 0.165 | 0.372 | 0.763 | 0.247 | 0.604 | 0.882 |
| Decile 8 | 0.367 | 0.650 | 0.777 | 0.855 | 0.942 | 0.050 | 0.111 | 0.276 | 0.277 | 0.552 | 0.834 | 0.153 | 0.361 | 0.737 | 0.266 | 0.501 | 0.809 | 0.154 | 0.491 | 0.693 |
| Decile 9 | 0.614 | 0.367 | 0.686 | 0.922 | 1.025 | 0.053 | 0.164 | 0.369 | 0.169 | 0.338 | 0.795 | 0.095 | 0.334 | 0.732 | 0.059 | 0.277 | 0.717 | 0.244 | 0.660 | 0.949 |
| Decile 10 | 0.585 | 0.415 | 0.622 | 0.768 | 0.843 | 0.039 | 0.229 | 0.401 | 0.313 | 0.599 | 0.818 | 0.205 | 0.460 | 0.816 | 0.103 | 0.354 | 0.821 | 0.232 | 0.616 | 0.814 |
| | | | | | | | | | **Bond Returns** | | | | | | | | | | | |
| 10 - Yrs Treasury | 0.707 | 0.249 | 0.848 | 0.962 | 1.099 | 0.024 | 0.068 | 0.112 | 0.209 | 0.449 | 0.837 | 0.117 | 0.456 | 0.759 | 0.481 | 0.674 | 0.808 | 0.157 | 0.364 | 0.482 |
| 5 - Yrs Treasury | 0.315 | 0.763 | 0.819 | 0.901 | 1.074 | 0.047 | 0.115 | 0.219 | 0.307 | 0.580 | 0.856 | 0.199 | 0.428 | 0.731 | 0.282 | 0.511 | 0.723 | 0.246 | 0.482 | 0.880 |
| Baa Corp Bonds (10-20 years) | 0.443 | 0.458 | 0.260 | 0.376 | 0.424 | 0.452 | 0.563 | 0.665 | 0.351 | 0.467 | 0.710 | 0.279 | 0.512 | 0.758 | 0.056 | 0.245 | 0.597 | 0.399 | 0.675 | 0.818 |

stocks exceed 0.5 with peaks in excess of 1 for a number of industries as well as medium-cap portfolios. However, the market factor does not explain most of the predictable variation in excess bond returns, when it is replaced in this leading role by the credit risk factor. As far as stocks are concerned, the next most important contributions come from unexpected inflation (especially for bond and selected industry portfolios) and to some extent, real consumption growth risk, although the heterogeneity across portfolios is pronounced (small capitalization stocks are particularly well explained by this factor). In the case of bond portfolios, most predictable variation is explained, after taking into credit risk exposures into account, by unexpected inflation, economy-wide market risks, and Cochrane and Piazzesi's specific factor; interestingly, the contribution of yield curve shocks to priced risk is limited.

### 5.2 Pricing Errors

We follow Geweke and Zhou (1996) and measure the closeness of the pricing approximation implied by (5), by computing at each time $t$ the average squared recursive pricing error across all the $N$ test assets/portfolios,

$$Q^2_{t,N} = \frac{1}{N}\left[\beta'_{0,t}\left(I_N - B_t\left(B'_tB_t\right)^{-1}B'_t\right)\beta_{0,t}\right] \qquad t = 1, ..., T, \tag{15}$$

where $\beta_{0,t}$ is the $N \times 1$ vector of intercepts, $B_t = (\iota_N, \beta_{1,t}, ..., \beta_{K,t})$ is a $N \times K$ matrix collecting vectors of time $t$ betas of all the assets/portfolios vs. each of the $K$ risk factors, and $\iota_N$ an $N$-dimensional vector of ones. These pricing errors are recursive because at each point in time they are obtained using only information available up to that point. Because our Gibbs sampling scheme allows to derive posteriors for all the objects that enter $\beta_{0,t}$ and $B_t$, we also compute the posterior density of the average pricing error statistic, as discussed in Geweke and Zhou (1996).

Table 6 reports the average monthly pricing errors, $Q^2_{t,N}$, for each of our models across different sub-samples. Using sub-samples wants to allow any instability in pricing performance to emerge and be adequately detected.With reference to the full-sample, the B-TVB-SV model yields both the lowest average pricing error (0.21% per month) and the lowest median posterior error (0.19%). Such statistics are practically between one-half and two-thirds those that one would obtain under a B-TVB homoskedastic model (0.41 and 0.35 percent, respectively).

Table 6: Average Pricing Errors

This table reports the average pricing errors for each of the models under investigation across different subsamples as well as in the full sample. *B-TVB-SV* stands for Bayesian time-varying betas, stochastic volatility model, while *B-TVB* and *B-TVP* are the dynamic Bayesian model restricted to have constant conditional volatility and random-walk betas, respectively. *F-MB* is the standard two-step procedure. The table reports the average (over time), the posterior standard deviation as well as the confidence interval at the 95% level.

| | **Average Pricing Errors** | | | | |
|---|---|---|---|---|---|
| | **Mean %** | **Std %** | **2.5 %** | **50 %** | **97.5 %** |
| **Panel A: Full-Sample** | | | | | |
| **B-TVB-SV** | **0.2108** | 0.0623 | 0.1363 | **0.1902** | 0.3231 |
| **B-TVB** | 0.4126 | 0.1588 | 0.2512 | 0.3459 | 0.7325 |
| **B-TVP** | 0.5401 | 0.1804 | 0.3113 | 0.5061 | 0.8126 |
| **F-MB** | 0.6303 | 0.0159 | 0.6107 | 0.6258 | 0.6633 |
| **Panel B: 1982:01 - 1999:01** | | | | | |
| **B-TVB-SV** | **0.1926** | 0.0431 | 0.1443 | **0.1851** | 0.2574 |
| **B-TVB** | 0.3935 | 0.0314 | 0.3511 | 0.3921 | 0.4521 |
| **B-TVP** | 0.5233 | 0.1653 | 0.3202 | 0.4759 | 0.8101 |
| **F-MB** | 0.6278 | 0.0151 | 0.6092 | 0.6238 | 0.6585 |
| **Panel C: 1999:01 - 2011:11** | | | | | |
| **B-TVB-SV** | **0.2624** | 0.0672 | 0.1454 | **0.2707** | 0.3525 |
| **B-TVB** | 0.4682 | 0.119 | 0.2806 | 0.4501 | 0.7068 |
| **B-TVP** | 0.6359 | 0.1993 | 0.3544 | 0.6456 | 0.9027 |
| **F-MB** | 0.6354 | 0.0162 | 0.6168 | 0.6321 | 0.6653 |
| **Panel D: 2007:01 - 2011:11** | | | | | |
| **B-TVB-SV** | **0.2891** | 0.0613 | 0.1673 | **0.2977** | 0.3952 |
| **B-TVB** | 0.5865 | 0.0906 | 0.4718 | 0.5713 | 0.7559 |
| **B-TVP** | 0.6397 | 0.1431 | 0.4029 | 0.6616 | 0.8168 |
| **F-MB** | 0.6523 | 0.0179 | 0.6164 | 0.6439 | 0.6799 |

Interestingly, the B-TVP model seems to fit the data well on the basis of the Bayes odds ratio in Table 4, but fails to price our test portfolios (it gives average and median posterior errors of 0.54 and 0.51 percent, respectively) as accurately as the homoskedastic B-TVB model does. The performance of the classical two-step F-MB scheme is poor, yielding average and median pricing errors of 0.63%. Moreover, B-TVB-SV consistently outperforms all other models in all sub-samples. Its advantage is always substantial in the sense that B-TVB-SV always cuts the average error of the second best model by at least 40%. The pricing errors tend to increase over our sample, especially when one compares the 1982-1998 with the 1999-2011 interval. However, there is no evidence of the errors during the Great Financial Crisis period (2007-2011) being systematically higher than in the overall 1999-2011 sub-sample.

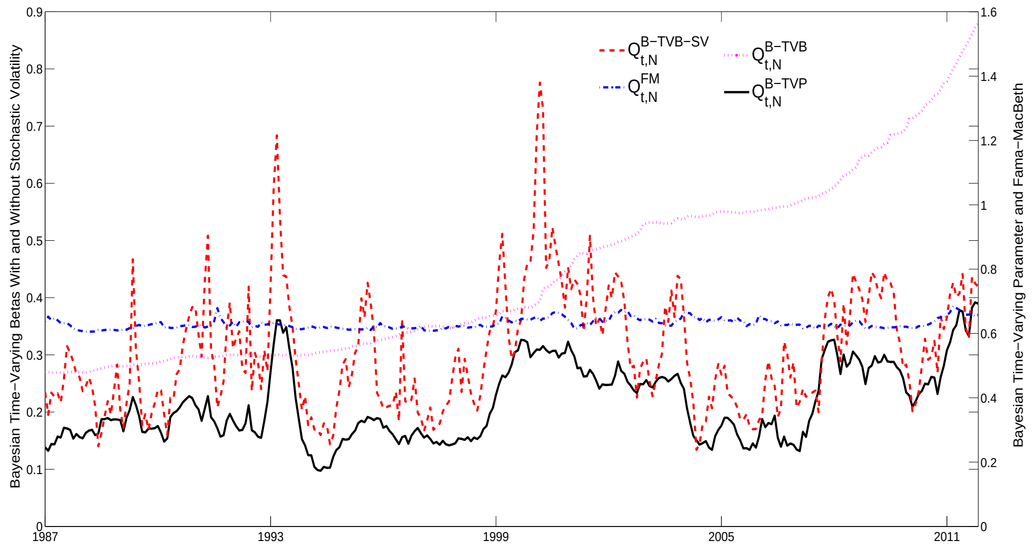
Figure 9 plots the time series of average pricing errors $Q^2_{t,N}$ for all the models (top panel) and only for the B-TVB-SV and the B-TVP cases, rescaling the errors from the former model (on the right axis) to better emphasize similarities and differences (bottom panel). The top panel shows that, apart from a short period in early 1993 (when B-TVB became competitive), B-TVB-SV gave uniformly lower average pricing errors than all other models. The F-MB scheme gives uniformly high but constant average errors. The B-TVP model gives a highly variable performance, with enormous spikes of mis-spricing around 1993, in 1999-2000, and during the financial crisis. The bottom panel of the figure shows that the dynamics of pricing errors under B-TVB-SV and B-TVP—both models including a stochastic volatility component—are not that dissimilar, in the sense that also errors from B-TVB-SV spike up in 1993, 1999-2001 and during 2011. However, the more parsimonious dynamics imposed by infrequent, large structural breaks under B-TVB-SV reduces the pricing errors also keeping the latter more stable across our sample.

## 6 Conclusion

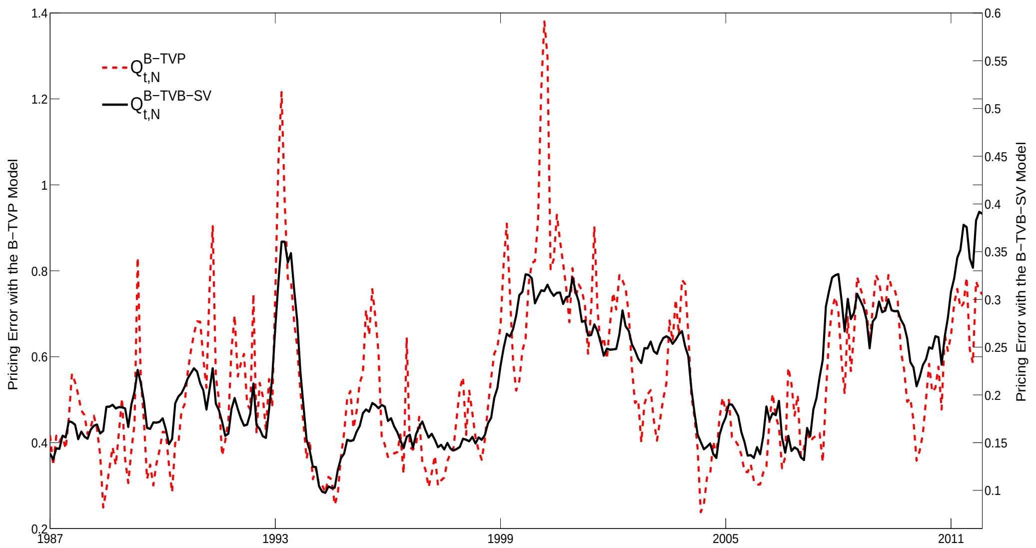
In this paper, we have proposed a new way to parameterize and estimate in state-space form a typical MFAPM with time-varying risk exposures and premia. This Bayesian state-space approach is based on a formal modelling of the latent process followed by risk exposures and idiosyncratic volatility capable to capture structural shifts in parameters. This method can also be interpreted as a novel way to overcome the two-pass approach advocated by Fama and

Figure 9: Average Pricing Errors

This figure reports the time series of the average pricing errors. The sample period is 1972:01 - 2011:12. The first ten years of data are used a training sample in order to calibrate the prior distribution for both latent states and parameters. Panel A reports the average pricing error across models. Panel B reports the rescaled values of the average pricing errors for the B-TVB-SV and the B-TVP models, respectively.




(a) Average Pricing Errors Across Models




(b) Average Pricing Errors for B-TVB-SV and B-TVP (Rescaled)

MacBeth (1973) and used in a substantive body of applied work in finance. Given a general B-TVB-SV framework, we have also considered special cases that are obtained by imposing restrictions, and in particular a B-TVP-SV model in which betas change continuously but in small amounts, and a homoskedastic B-TVB model in which volatility is constant.

Our application to monthly, 1972-2011 U.S. stock and bond returns shows that the two-stage approach yields results that are not always reasonable. For instance, very few risk factors appear to be priced and, when they are, they carry the wrong sign. Moreover, the fit provided by the standard two-step approach is poor. On the contrary, the empirical implications of a Bayesian state-space implementation are plausible and there are indications that the model is consistent with the data. For instance, most portfolios do not appear to have been grossly mispriced and a few risk premia are precisely estimated with a plausible sign. Market, liquidity, and industrial production (real output) growth risks are significantly priced. This confirms the early evidence in Burmeister and McElroy (1988) that appropriate econometric methods reveal a strong explanatory power of macroeconomic factors in addition to that provided by the plain vanilla, CAPM-style market portfolio. Bayes odds ratios and marginal likelihood comparisons indicate that the B-TVB-SV outperforms both the two-step F-MB and the homoskedastic B-TVB models. The heteroskedastic B-TVP appears to be closer to the full-scale B-TVB-SV one. However, an analysis of the average pricing errors shows that large but infrequent breaks in factor exposures are considerably more successful. Finally, the finding that the heteroskedastic B-TVB models ranks second below B-TVB-SV is a powerful indication of the importance to explicitly model stochastic volatility when implementing multi-factor asset pricing models.

## Supplementary Material

**Online Appendix:** In the online appendix, we provide additional detail regarding our methodology and some additional results. In this appendix, Appendix A describes in detail the Gibbs sampler used for estimating the model. Convergence properties for our MCMC approach can be found in Appendix B, whereas Appendix C reports on a prior sensitivity analysis for our framework. Appendix D investigates the impact of different specification of stochastic volatility on in-sample posterior inference. Finally, Appendix E details the variance decomposition test used to assess the economic performances of the model.

## Acknowledgments

We are grateful to Carlos M. Carvalho, Fabio Trojani (discussant) and to seminar participants at the University of St.Gallen, Bocconi University, Norges Bank, the $1_{st}$ Vienna workshop on high-dimensional time series in macroeconomics and finance, the $5_{th}$ Italian Congress of Econometrics and Empirical Economics, the SIS 2013 Statistical Conference, the $6_{th}$ International Conference on Computational and Financial Econometrics and the 2014 European Financial Association meetings for helpful comments and suggestions. Ravazzolo is also affiliated with the BI Norwegian Business School. The views expressed in this paper are our own and do not necessarily reflect those of Norges Bank.

## References

Ang, A. and Chen, J. (2007). Capm over the long-run: 1926-2001. *Journal of Empirical Finance*, (14):1–40.

Bekaert, G., Hodrick, R., and Zhang, X. (2012). Aggregate idiosyncratic volatility. *Journal of Financial and Quantitative Analysis*, 47:1155–1185.

Bossaerts, P. and Green, R. (1989). A general equilibrium model of changing risk premia: Theory and tests. *Review of Financial Studies*, 2:467–493.

Burmeister, E. and McElroy, M. (1988). Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory. *Journal of Finance*, 43:721–733.

Campbell, J., Lettau, M., Malkiel, B., and Xu, Y. (2001). Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk. *Journal of Finance*, 56:1–43.

Carter, C. and Kohn, R. (1994). On gibbs sampling for state-space models. *Biometrika*, (81):541–553.

Chan, L., Karceski, J., and Lakonishok, J. (1998). The risk and return from factors. *Journal of Financial and Quantitative Analysis*, 33:159–188.

Chen, N., Roll, R., and Ross, S. (1986). Economic forces and the stock market. *Journa of Business*, 59:383–403.

Chib, S. (1995). Marginal likelihood from the gibbs output. *Journal of the American Statistical Association*, 432:1313–1321.

Clark, T. and Ravazzolo, F. (2014). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, Forthcoming.

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

Cochrane, J. and Piazzesi, M. (2005). Bond risk premia. *American Economic Review*, (94):138–160.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance*, (57):369–403.

Eisenstat, E. and Strachan, R. (2014). Modelling inflation volatility. *CAMA Working Paper 21/2014*.

Fama, E. and MacBeth, J. (1973). Risk, return and equilibrium. *Journal of Political Economy*, (81):607–636.

Ferson, W. and Harvey, C. (1991). The variation of economic risk premiums. *Journal of Political Economy*, 99:385–415.

Flannery, M. and Protopapadakis, A. (2002). Macroeconomic factors do influence aggregate stock returns. *Review of Financial Studies*, 15:751–782.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15:183–202.

Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions*, 6:721–741.

George, E. and McCulloch, R. (1993). Variable selection via gibbs sampling. *Journal of the American Statistical Association*, (88):881–889.

Gerlach, R., Carter, C., and Kohn, R. (2000). Efficient bayesian inference for dynamic mixture models. *Journal of the American Statistical Association*, (95):819–828.

Geweke, J. and Amisano, G. (2010). Comparing and evaluating bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26:216–230.

Geweke, J. and Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies*, 9:557–587.

Giordani, P., K. R. and van Dijk, D. (2007). A unified approach to nonlinearity, outliers and structural breaks. *Journal of Econometrics*, 137:112–133.

Giordani, P. and Kohn, R. (2008). Efficient bayesian inference for multiple change-point and mixture innovation models. *Journal of Business and Economic Statistics*, 26:66–77.

Giordani, P. and Villani, M. (2010). Forecasting macroeconomic time series with locally adaptive signal extraction. *International Journal of Forecasting*, 26:312–325.

Groen, J., Paap, R., and Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business and Economic Statistics*, 31:29–44.

Gungor, S. and Luger, R. (2013). Testing linear factor pricing models with large cross sections: A distribution-free approach. *Journal of Business and Economic Statistics*, (31):66–77.

Hansen, P. and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a garch(1,1)? *Journal of Applied Econometrics*, 20:873–889.

Harvey, C. (2001). The specification of conditional expectations. *Journal of Empricial Finance*, (8):573–638.

Jostova, G. and Philipov, A. (2005). Bayesian analysis of stochastic betas. *Journal of Financial and Quantitative Analysis*, 40:747–778.

Karolyi, G. and Sanders, A. (1998). The variation of economic risk premiums in real estate returns. *Journal of Real Estate Finance and Economics*, 17:245–262.

Kass, R. and Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, 430:773–795.

Kim, S., Shepard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *Review of Economic Studies*, 65:361–393.

Koop, G. and Potter, S. (2007). Estimation and forecasting in models with multiple breaks. *Review of Economic Studies*, 74:763–789.

Kramer, C. (1994). Macroeconomic seasonality and the january effect. *Journal of Finance*, 49:1883–1891.

Lamont, O. (2001). Economic tracking portfolios. *Journal of Econometrics*, 105:161–184.

Lettau, M. and Ludvigson, S. (2001). Resurrecting the (c)capm: A cross-sectional test when risk premia are time varying. *Journal of Political Economy*, (109):1238–1287.

Maheu, J. and Gordon, S. (2008). Learning, forecasting and structural breaks. *Journal of Applied Econometrics*, 23:553–583.

Maheu, J. and McCurdy, T. (2009). How useful are historical data for forecasting the long-run equity return distribution? *Journal of Business and Economic Statistics*, 27:95–112.

McCulloch, R. and Roley, V. (1993). Stock prices, news, and business conditions. *Review of Financial Studies*, 6:683–707.

McCulloch, R. and Rossi, P. (1991). Posterior, predictive and utility based approaches to testing arbitrage pricing theory. *Journal of Financial Economics*, 28:7–38.

Merton, R. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41:867–887.

Miazhynskaia, T., Frühwirth-Schnatter, S., and Dorffner, G. (2006). Bayesian testing for non-linearity in volatility modeling. *Computational Statistics and Data Analysis*, 51:2029–2042.

Mitchell, T. and Beauchamp, J. (1988). Bayesian variable selection in linear regression. *Journal of the American Statistical Association*, 83:1023–1032.

Nardari, F. and Scruggs, J. (2007). Bayesian analysis of linear factor models with latent factors, multivariate stochastic volatility, and apt pricing restrictions. *Journal of Financial and Quantitative Analysis*, 42(4):857–891.

Omori, Y., Chib, S., Shepard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140:425–449.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, 25:221–247.

Pastor, L. and Stambaugh, R., F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, (111):642–685.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22:435–480.

Primiceri, G. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72:821–852.

Ravazzolo, F., Paap, R., van Dijk, D., and Franses, P. (2007). *Bayesian Model Averaging in the Presence of Structural Breaks*. Elsevier, frontiers of economics and globalization edition.

Shanken, J. (1992). On the estimation of beta pricing models. *Review of Financial Studies*, (5):1–34.

Shanken, J. and Weinstein, M. (2006). Economic forces and the stock market revisited. *Journal of Empirical Finance*, (13):129–144.

Shiller, R. (1979). The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy*, 87:1190–1219.

Singleton, K. (2006). *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. Princeton University Press, Princeton: NJ.

Tanner, A. and Wong, W. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–540.

Thaler, R. (1987). Anomalies: The january effect. *Journal of Economic Perspectives*, 1:197–201.

West, M. and Harrison, J. (1997). *Bayesian forecasting and dynamics models*. Springer.

# Appendix

## A The Gibbs Sampling Algorithm

In this section we derive the full conditional posterior distributions of the latent variables and the model parameters discussed in Section 2 of the main text. Before we describe in detail the different steps of sampler, we need to define the densities that make up the joint density of the data and the latent variables (12). By considering the no-arbitrage restriction such that $\beta_{i0,t} \simeq \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t-1}$, the likelihood given states and parameters can be written from (4)-(5) as

$$p(r_{it}|F_t, \beta_{it}, \sigma_{it}^2) = \frac{1}{\sqrt{2\pi\sigma_{i,t}^2}} \exp\left(-\frac{\left(r_{it} - \beta_{i0,t} - \sum_{j=1}^{K} \beta_{ij,t}F_{j,t}\right)^2}{2\sigma_{i,t}^2}\right)$$

from (6)-(8) the densities of the latent states can be written as

$$p(\beta_{i,t}|\beta_{i,t-1}, \kappa_{i,t}, q_i^2) = \prod_{j=0}^{K} \left(\frac{1}{\sqrt{2\pi q_{ij}^2}} \exp\left(-\frac{(\beta_{ij,t} - \beta_{ij,t-1})^2}{2q_{ij}^2}\right)\right)^{\kappa_{ij,t}} (\beta_{ij,t} - \beta_{ij,t-1})^{1-\kappa_{ij,t}}$$

$$p(\ln\sigma_{it}^2|\ln\sigma_{it-1}^2, \kappa_{iv,t}, q_{iv}^2) = \left(\frac{1}{\sqrt{2\pi q_{iv}^2}} \exp\left(-\frac{\left(\ln\sigma_{i,t}^2 - \ln\sigma_{i,t-1}^2\right)^2}{2q_{iv}^2}\right)\right)^{\kappa_{iv,t}} \left(\ln\sigma_{i,t}^2 - \ln\sigma_{i,t-1}^2\right)^{1-\kappa_{iv,t}} \quad \text{(A.1)}$$

The densities for $\beta_{i,t}$ and $\ln\sigma_{it}^2$ each consist of two parts. First one where breaks occurs and these are drawn from their corresponding distributions. The second component is the case of no break which results in a degenerate distribution of either the $\beta_{ij,t}$ or $\ln\sigma_{it}^2$. Note the latter case my be also represented as a Dirac delta function. For the ease of exposition we summarize the Gibbs sampler for the $i_{th}$ asset.

### A.1 *Step 1. Sampling*$K_\beta$.

The structural breaks in the conditional dynamics of the factor loadings $B$, measured by the latent binary state $\kappa_{jt}$, are drawn using the algorithm of Gerlach et al. (2000). This algorithm increases the efficiency of the sampling procedure since allows to generate $\kappa_{jt}$, without conditioning on the relative regression parameters $\beta_{jt}$. The conditional posterior density for $\kappa_{jt}$, $t = 1, ..., T, j = 0, ..., K$, is defined as

$$\begin{aligned} p\left(\kappa_{0t,...,}\kappa_{Kt}|\mathcal{K}_{\beta[-t]}, \mathcal{K}_\sigma, \Sigma, \theta, R, F\right) &\propto p\left(R|\mathcal{K}_{\beta t}, \mathcal{K}_\sigma, \Sigma, \theta, F\right) p\left(\kappa_{0t,...,}\kappa_{Kt}|\mathcal{K}_{\beta[-t]}, \mathcal{K}_\sigma, \Sigma, \theta, F\right) \\ &\propto p\left(r_{t+1}, ..., r_T|r_1, ..., r_t, \mathcal{K}_{\beta t}, \mathcal{K}_\sigma, \Sigma, \theta, F\right) p\left(r_t|r_1, ..., r_{t-1}, \kappa_{0t,...,}\kappa_{Kt}, \mathcal{K}_\sigma, \Sigma, \theta, F\right) \\ &\quad p\left(\kappa_{0t,...,}\kappa_{Kt}|\mathcal{K}_{\beta[-t]}, \mathcal{K}_\sigma, \Sigma, \theta, F\right) \end{aligned} \quad \text{(A.2)}$$

where $\mathcal{K}_{\beta[-t]} = \left\{\{\kappa_{js}\}_{j=0}^{K}\right\}_{s=1, s\neq t}^{T}$. We assume that each of the $\kappa_{js}$ breaks are independent from each other such that the joint density is defined is defined as $\prod_{j=0}^{K} \pi_{ij}^{\kappa_{jt}} (1 - \pi_{ij})^{1-\kappa_{jt}}$.

The remaining densities $p\left(r_{t+1}, ..., r_T|r_1, ..., r_t, \mathcal{K}_{\beta t}, \mathcal{K}_\sigma, \Sigma, \theta, F\right)$ and $p\left(r_t|r_1, ..., r_{t-1}, \kappa_{0t,...,}\kappa_{Kt}, \mathcal{K}_\sigma, \Sigma, \theta, F\right)$ are evaluated as in Gerlach et al. (2000). Notice that, since $\kappa_{jt}$ is a binary state the integrating constant is easily evaluated.

### A.2 *Step 2. Sampling the Factor Loadings* B.

The full conditional posterior density for the time-varying factor loadings is computed using a standard forward filtering backward sampling as in Carter and Kohn (1994). For each of the $i = 1, ..., N$ assets, the prior distribution of the $\beta_{i0}, ..., \beta_{iK}$ loadings is a multivariate normal with the location parameters corresponding to the OLS parameter estimates and a covariance structure which is diagonal and defined by the variances of the OLS estimates. The initial prior are sequentially updated via the Kalman Filtering recursion, then the parameters are drawn from the posterior distribution which is generated by a standard backward recursion (see Frühwirth-Schnatter 1994, Carter and Kohn 1994, and West and Harrison 1997).

## A.3 *Step 3 and 4. Sampling the Breaks and the Values of the Idiosyncratic Volatility.*

In order to draw the structural breaks $\mathcal{K}_\sigma$ and the idiosyncratic volatilities $S$ we follow a similar approach as above. The stochastic breaks $\mathcal{K}_\sigma$ are drawn by using the Gerlach et al. (2000) algorithm. The conditional variances $\ln \sigma^2_{it}$, does not show a linear structure even though still preserving the standard properties of state space models. The model is rewritten as

$$\ln \left( r_{i,t} - \beta_{0t} - \sum_{j=1}^{K} \beta_{ijt} F_{jt} \right)^2 = \ln \sigma^2_{it} + u_t$$
$$\ln \sigma^2_{it} = \ln \sigma^2_{it-1} + \kappa_{\nu it} \nu_{it} \tag{A.3}$$

where $u_t = \ln \varepsilon^2_t$ has a $\ln \chi^2(1)$. Here we follow Omori et al. (2007) and approximate the $\ln \chi^2(1)$ distribution with a finite mixture of ten normal distributions, such that the density of $u_t$ is given by

$$p(u_t) = \sum_{l=1}^{10} \varphi_l \frac{1}{\sqrt{\varpi^2_l 2\pi}} \exp\left( -\frac{(u_t - \mu_l)^2}{2\varpi_l} \right) \tag{A.4}$$

with $\sum_{l=1}^{10} \varphi_l = 1$. The appropriate values for $\mu_l, \varphi_l$ and $\varpi^2_l$ can be found in Omori et al. (2007). Mechanically in each step of the Gibbs Samplers we simulate at each time $t$ a component of the mixture. Now, given the mixture component we can apply the standard Kalman filter method, such that $\mathcal{K}_\sigma$ and $\Sigma$ can be sampled in a similar way as $\mathcal{K}_\beta$ and $B$ in the first and second step. The initial prior of the log idiosyncratic volatility $\ln \sigma^2_0$ is normal with mean -1 and conditional variance equal to 0.1.

## A.4 *Step 5a. Sampling the Risk Premia at Time* $t$.

The equilibrium restriction in (2) simplify at each time $t$ to a multi-variate linear regression of the $N$ excess returns $r = (r_{1,t}, r_{2,t}, ..., r_{N,t})'$, onto a constant term and past betas $X = (\iota_N, \beta_{1,t-1}, \beta_{2,t-1}, ..., \beta_{K,t-1})$

$$r = X\lambda + e \qquad \text{with} \qquad e \sim N(0, \tau^2 I_N) \tag{A.5}$$

where $\beta_{i,t-1} = (\beta_{1i,t}, \beta_{2i,t}, ..., \beta_{Ni,t})'$. Note here we avoid the time $t$ dependence of regressors for the ease of exposition. We consider independent conjugate priors

$$\lambda \sim MN(\underline{\lambda}, \underline{V}) \qquad \tau^2 \sim IG-2(\overline{\psi}_0, \Psi_0) \tag{A.6}$$

Posterior updating in the Gibbs sampler evolves as

$$\lambda | X, r \sim MN(\overline{\lambda}, \overline{V}) \qquad \tau^2 | X, r \sim IG-2(\overline{\psi}, \Psi) \tag{A.7}$$

with

$$\overline{\lambda} = \overline{V}\left( \underline{V}^{-1} \underline{\lambda} + \tau^2 X' r \right) \qquad \text{and} \qquad \overline{V} = \left( \underline{V}^{-1} + \tau^{-2} X' X \right) \tag{A.8}$$

while the posterior hyper-parameters for conditional volatility are defined as

$$\overline{\psi} = \overline{\psi}_0 + N \qquad \text{and} \qquad \Psi = \Psi_0 + ee' \tag{A.9}$$

## A.5 *Step 5b. Sampling the Stochastic Breaks Probabilities.*

The full conditional posterior densities for the breaks probabilities $\pi = (\pi_{i1,...,}\pi_{iK})$ is given by

$$p\left(\pi|q^2, B, \mathbf{\Sigma}, \mathcal{K}_\beta, R, F\right) \propto \prod_{j=0}^{K} \pi_{ij}^{a_{ij}-1}\left(1-\pi_{ij}\right)^{b_{ij}-1} \prod_{t=1}^{T} \pi_{ij}^{\kappa_{ijt}}\left(1-\pi_{ij}\right)^{1-\kappa_{ijt}} \tag{A.10}$$

and hence the individual $\pi_{ij}$ parameter can be sampled from a Beta distribution with shape parameters $a_{ij} + \sum_{t=1}^{T}\kappa_{ijt}$ and $b_{ij} + \sum_{t=1}^{T}(1-\kappa_{ijt})$ for $j = 0, ..., K$. Likewise the full conditional posterior distribution for the breaks probabilities in the idiosyncratic volatilities $\pi_\nu$ is given by

$$p\left(\pi_\nu|q^2, B, \mathbf{\Sigma}, \mathcal{K}_\sigma, R, F\right) \propto \pi_{i\nu}^{a_{i\nu}-1}\left(1-\pi_{i\nu}\right)^{b_{i\nu}-1} \prod_{t=1}^{T} \pi_{i\nu}^{\kappa_{i\nu t}}\left(1-\pi_{i\nu}\right)^{1-\kappa_{i\nu t}}$$

such that the individual $\pi_{i\nu}$ can be sampled from a Beta distribution with shape parameters $a_{i\nu} + \sum_{t=1}^{T}\kappa_{i\nu t}$ and $b_{i\nu} + \sum_{t=1}^{T}(1-\kappa_{i\nu t})$ for $i = 1, ..., N$.

## A.6 *Step 5c. Sampling the Conditional Variance of the States.*

The prior distributions for the conditional volatilities of the factor loadings $\beta_{ijt}$ for $j = 0, ..., K$ are inverse-gamma

$$p\left(q_{ij}^2|\pi, B, \mathbf{\Sigma}, \mathcal{K}_\beta, \mathcal{K}_\sigma, R, F\right) \propto q_{ij}^{-\nu_{ij}} \exp\left(-\frac{\delta_{ij}}{2q_{ij}^2}\right) \prod_{t=1}^{T}\left(\frac{1}{q_{ij}}\exp\left(\frac{-\left(\beta_{ijt-}\beta_{ijt-1}\right)^2}{2q_{ij}^2}\right)\right)^{\kappa_{ijt}} \tag{A.11}$$

hence $q_{ij}^2$ is sampled from an inverse-gamma distribution with scale parameter $\nu_{ij} + \sum_{t=1}^{T}\kappa_{ijt}\left(\beta_{ijt-}\beta_{ijt-1}\right)^2$ and degrees of freedom equal to $\nu_{ij} + \sum_{t=1}^{T}\kappa_{ijt}$. Likewise the full conditional of the variance for the idiosyncratic log volatility $q_{i\nu}^2$ is defined as

$$p\left(q_{i\nu}^2|\pi, B, \mathbf{\Sigma}, \mathcal{K}_\beta, \mathcal{K}_\sigma, R, F\right) \propto q_{i\nu}^{-\nu_{i\nu}} \exp\left(-\frac{\delta_{i\nu}}{2q_{i\nu}^2}\right) \prod_{t=1}^{T}\left(\frac{1}{q_{i\nu}}\exp\left(\frac{-\left(\ln\sigma_{it}^2 - \ln\sigma_{it-1}^2\right)^2}{2q_{i\nu}^2}\right)\right)^{\kappa_{i\nu t}} \tag{A.12}$$

such that $q_{i\nu}^2$ is sampled from an inverted Gamma distribution with scale parameter $\nu_{i\nu} + \sum_{t=1}^{T}\kappa_{i\nu t}\left(\ln\sigma_{it}^2 - \ln\sigma_{it-1}^2\right)^2$ and degrees of freedom equal to $\nu_{i\nu} + \sum_{t=1}^{T}\kappa_{i\nu t}$.

# B MCMC Convergence Analysis

We report the results of a convergence analysis of the MCMC sampler for the B-TVB-SV model outlined in Section 3 and Appendix A. The convergence analysis involves computing a set of inefficiency factors and t-tests for equality of the means across subsamples of the MCMC chain. (see Geweke 1992, Primiceri 2005 Justiniano and Primiceri 2008, Clark and Davig 2011 and Groen et al. 2013).

For each individual parameter and latent variable, the inefficiency factor answer the question "How much information do we actually have about parameters?", and is measured as $(1+2\sum_{f=1}^{\infty}\rho_f)$, where $\rho_f$ is the $f_{th}$ order auto-correlation of the chain of draws. This inefficiency factor equals the variance of the mean of the posterior draws from the MCMC sampler, divided by the variance of the mean assuming independent draws. Then, if we require that the variance of the mean of the MCMC posterior draws should be limited to be at most 1% of the variation due to the data (measured by the posterior variance), the inefficiency factor provides an indication of the minimum number of MCMC draws to achieve this, see Kim et al. (1998). If there are some correlation between successive samples, then we might expect that our sample has not revealed as much information of the posterior distribution of our parameter as we could have gotten if the samples draws were independent. When estimating these inefficiency factors, we use the Bartlett kernel as in Newey and West (1987), with a bandwidth set to 4% of the sample of draws. The inefficiency factor is computed for all the model parameters and applied on a range of choices for the total number of posterior draws as well as burn-in period lengths and thinning for the B-TVB-SV specification. Based on this comparison we felt most comfortable that with the number of posterior draws set equal to 10000 with a burn-in period of 2000 draws and thinning value of 2, yielding 10000 retained posterior draws, our MCMC sampler would perform satisfactorily.

Tables B.1 provide a summary of the results showing that, for most parameters and latent variables, our MCMC sampler is very efficient and that it requires far less than 5000 retained posterior draws to be able to do a reasonably accurate inferential analysis. In case of the time-invariant parameters $\mathbf{Q}$ and $\pi$, with likely values in the 2.3-4.2 range, our sampler is less efficient. Nonetheless, the corresponding inefficiency factors suggest on average a minimum number of draws of less than 4000 to achieve an accurate analysis of these parameters.

Table B.1: Summary of Inefficiency Factors

The table summarizes the inefficiency factors, for the posterior values of the model parameters, estimated over the sample period 1972:01 - 2011:12. The estimated inefficiency factors are based on the Bartlett kernel as in Newey and West (1987) with a bandwidth equal to 4% of the 10000 retained draws.

| | Inefficiency Factor | | | | | | |
|---|---|---|---|---|---|---|---|
| | Parameters | Mean | Median | Min | Max | 5% | 95% |
| $\mathbf{B}$ | 82800 | 2.9081 | 2.9222 | 2.6001 | 3.5031 | 2.6842 | 3.2091 |
| $\mathcal{K}_\beta, \mathcal{K}_\sigma$ | 91080 | 2.7886 | 2.8157 | 2.0096 | 4.0311 | 2.3567 | 3.4231 |
| $\mathbf{\Sigma}$ | 8280 | 2.8121 | 2.8321 | 2.0897 | 3.9421 | 2.4016 | 3.4072 |
| $\mathbf{Q}$ | 253 | 2.9318 | 2.9118 | 2.3314 | 3.8921 | 2.3414 | 3.8532 |
| $\pi$ | 253 | 3.3478 | 3.3405 | 2.6652 | 4.2307 | 2.6668 | 4.2209 |

We also compute the p-value of the Geweke (1992) t-test for the null hypothesis of equality of the means computed with the first 20 percent and last 40 percent of the sample of retained draws. For this particular convergence diagnostic test we compute the variances of the respective means using the Newey and West (1987) heteroskedasticity and autocorrelation robust variance estimator with a bandwidth set to 4% of the utilized sample sizes. Such convergence statistics is still computed for the complete B-TVB-SV specification estimated over the sample period 1972:01 - 2011:12. Table B.2 shows the results. The convergence diagnostic tests in Table

Table B.2: Summary of Convergence Diagnostics

The table summarizes the convergence results, for the posterior values of the model parameters, estimated over the sample period 1972:01 - 2011:12. For each of these, we compute the p-value of the Geweke (1992) t-test for the null hypothesis of equality of the means computed for the first 20% and the last 40% of the retained 10000 draws. The variances of the means are estimated with the Newey and West (1987) variance estimator using a bandwidth of 4% of the respective sample sizes.

| | Summary of Convergence Diagnostics | | |
|---|---|---|---|
| | Parameters | 5% Reject Rate | 10% Reject Rate |
| $\mathbf{B}$ | 82800 | 0.0102 | 0.0347 |
| $\mathcal{K}_\beta, \mathcal{K}_\sigma$ | 91080 | 0.0133 | 0.0400 |
| $\mathbf{\Sigma}$ | 8280 | 0.0108 | 0.0317 |
| $\mathbf{Q}$ | 253 | 0.0000 | 0.0000 |
| $\pi$ | 253 | 0.0000 | 0.0000 |

B.2 confirm the efficiency of the MCMC sampler we propose. For example, in the case of the $\mathbf{B}$ parameters the null hypothesis of equal means across sub-samples of the retained draws is hardly ever rejected at the 5% confidence interval. Thus, inference in our factor model appears to be reasonably accurate when we base posterior inference on 10000 draws with a burn-in of 2000 and thin value of 2. Such a choice of the number of draws keeps the computational burden relatively low, at the benefit of inference precision as shown in Table B.1 and Table B.2.

# C Prior Sensitivity Analysis

We investigate in this section the influence of different prior specifications on posterior results. In particular we discuss prior sensitivity for both the expected occurrence probability and expected size of a break for betas and idiosyncratic risks. First, we run a simulation example and directly test how posterior estimates reacts to different prior specifications. Second, we estimate the B-TVB-SV model on the original dataset by using different priors specifications. The goal of both exercises is to assess how the model instability/dynamics implied by posterior estimates is driven by priors on break sizes and probabilities.

## C.1 Simulation Example

The first step of the prior sensitivity analysis is based on a simulation example. We base our results on the following data generating process [DGP]

$$y_t = \beta_{0,t} + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \sigma_t\epsilon_t, \qquad \text{for} \qquad t = 1, ..., 200$$

with $\epsilon_t \sim NID(0,1)$ and $x_{j,t} \sim NID(0,1)$ for $j = 1, ..., 3$. We simulate discrete breaks both in the betas and idiosyncratic risks. The intercept is set to $\beta_{0,t} = 0$ for $t = 1, ..., 200$, meaning we simulate a factor model where there is no pricing error in the DGP. For the first regressor we take as parameters $\beta_{1,t} = 0.4$ for $t = 1, ..., 80$, $\beta_{1,t} = 0.9$ for $t = 81, ..., 160$, $\beta_{1,t} = 0.1$ and $t = 161, ..., 200$. For the second regressor we have $\beta_{2,t} = 0.2$ for $t = 1, ..., 60$, $\beta_{2,t} = 0.5$ for $t = 61, ..., 120$, and $\beta_{2,t} = 0$ for $t = 121, ..., 200$. Furthermore, $\beta_{3,t} = -0.2$ for $t = 1, ..., 60$, $\beta_{3,t} = -0.5$ for $t = 61, ..., 150$, and $\beta_{3,t} = -0.1$ for $t = 151, ..., 200$. For the (log of) idiosyncratic volatility we assume that $\ln \sigma_t^2 = -3$ for $t = 1, ..., 60$, $\ln \sigma_t^2 = -1.5$ for $t = 61, ..., 140$, and $\ln \sigma_t^2 = -2$ for $t = 141, ..., 200$. Hence we allow for breaks in the parameters at different points in time but we also include breaks which occur at the same time.

We apply our Bayesian estimation framework with structural breaks outlined in Section 3 with $M = 10000$ posterior draws (burn-in of 2000 draws and thin of 2), and different prior settings to investigate the sensitivity of posterior results. As a base case we assume the hyper-parameters outlined in the main text. We set $a_j = 3.2$, $b_j = 60$ and $\gamma_j = 0.5$, $\delta_j = 100$ for $j = 1, 2, 3$ which implies *a priori* a relatively low break probability of having a break with a moderate expected size. As far as the (log of) idiosyncratic risk is concerned, we assume for the base case a low probability of having a break with $a_\nu = 1$, $b_\nu = 99$ while the size of breaks implies $\gamma_j = 0.2$,$\delta_j = 50$. We report in Figure C.1 the posterior estimates of $\beta_{i,t}$ for $i = 1, 2, 3$ and $\ln \sigma_t^2$ together with the corresponding *true* parameters. The results from this figure show that our approach is quite accurate in estimating both the timing and the size of the breaks, where the estimates of $\beta_{i,t}$ for $i = 1, 2, 3$ are more volatile due to our prior choice for the hyper-parameters of the inverse-gamma distributed size of the breaks. As we would expected, conditional volatility of the estimates sensibly increases around the occurrence of breaks in the DGP.

In our prior sensitivity assessment, we consider several alternative prior specifications where increase the prior probability of a break and decrease or increase the expected size of the breaks both across betas and idiosyncratic volatility. In total we consider 12 different prior specifications. A moderately larger probability of a break than in the base case means that we divide $b_j$ by 5. A more extreme break probability is obtained by dividing $b_j$ by 10. Yet, a higher (lower) expected prior break size is obtained by multiplying (dividing) $\gamma_j$ and $\nu_j$ by 5. As far as the conditional volatility is concerned, we increase the probability of a break by dividing $b_\nu$ by 10 and 20, respectively. Table C.3 summarizes the different prior settings. To summarize, by considering $b_j = 6, 12$ we increase the prior expected probability of observing a break in the betas (idiosyncratic risk) to 20% and 35% (10% and 17%) respectively.

Figure C.1: Posterior Estimates of Time-Varying Parameters for the Base case Priors

This figure plots the posterior distributions of the parameters $\beta_{i,t}$ for $i = 1, 2, 3$ and $\ln \sigma_t^2$ together with the corresponding values implied by the DGP. The blue dashed line reports the median estimates of the parameters. The red dashed lines denote the 20th and 80th percentiles of the posterior distribution. The black solid line displays the values of the data generating process.

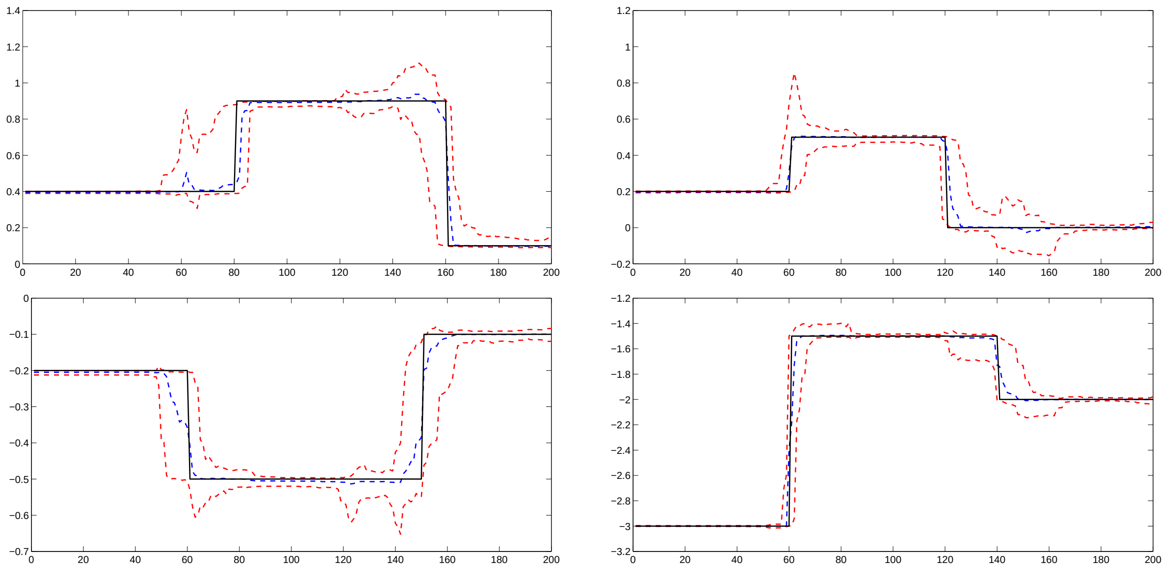


Table C.3: Summary of Prior Settings for Different Cases

The table summarizes the different prior settings we used to run the prior sensitivity analysis.

| Break | Exp | Prior Betas | | | | Prior Variance | | | |
|---|---|---|---|---|---|---|---|---|---|
| Probability | Size | $a_j$ | $b_j$ | $\gamma_j$ | $\delta_j$ | $a_\nu$ | $b_\nu$ | $\gamma_\nu$ | $\delta_\nu$ |
| Base | | 3.2 | 60 | 0.5 | 100 | 1 | 99 | 0.2 | 50 |
| Large | Small | 3.2 | 12 | 0.1 | 20 | 1 | 10 | 0.04 | 10 |
| Large | Large | 3.2 | 12 | 2.5 | 500 | 1 | 10 | 1 | 250 |
| Higher | Small | 3.2 | 6 | 0.1 | 20 | 1 | 5 | 0.04 | 10 |
| Higher | Large | 3.2 | 6 | 2.5 | 500 | 1 | 5 | 1 | 250 |

Figure C.2 reports the posterior estimates of $\beta_{i,t}$ for $i = 1, 2, 3$ and $\ln \sigma_t^2$ by increasing the prior probability of having a break ($a_j = 3.2$, $b_j = 12$, $a_\nu = 1$ and $b_\nu = 10$), as well as rising the prior average size of the breaks ($\gamma_j = 2.5$, $\delta_j = 500$, $\gamma_\nu = 1$ and $\delta_\nu = 250$). The figure makes it clear that posterior medians of the parameter are now quite off in terms of the timing of the breaks, with a large uncertainty for the posterior estimates of $\beta_{1:3,t}$ as well as $\ln \sigma_t^2$. Interestingly, higher uncertainty is more evident for the betas than for the log of conditional variance. Indeed, although posterior median estimates of $\ln \sigma_t^2$ are fairly off from capturing the second break point, the corresponding confidence intervals are still relatively tight. As we would expect, by imposing *a priori* a higher instability in the parameters of the model, the corresponding credibility intervals tend to increase. Figure C.3 shows the posterior estimates of the model parameters by assuming an even larger prior break probability ($a_j = 3.2$, $b_j = 6$, $a_\nu = 1$ and $b_\nu = 5$), while still keeping the ex-ante average break size as before ($\gamma_j = 2.5$, $\delta_j = 500$, $\gamma_\nu = 1$ and $\delta_\nu = 250$). The Figure shows that a higher expected probability and size of a break may lead to much more uncertain posterior estimates. From Figure C.3-C.2, however, is hard to say if less precise estimates comes from a higher expected probability, rather than a higher expected size of a break.

## Figure C.2: Posterior Estimates of the Time-Varying Parameters: Large prior-break probabilities and large prior break size

This figure plots the posterior distributions of of the parameters $\beta_{i,t}$ for $i = 1, 2, 3$ and $\ln \sigma_t^2$ together with the corresponding values implied by the DGP. The blue dashed line reports the median estimates of the parameters. The red dashed lines denote the 20th and 80th percentiles of the posterior distributions. The black solid line displays the values used to define the data generating process. Posterior results are now based on a larger prior-break probabilities for both betas and the (log of) idiosyncratic volatility ($a_j = 3.2$, $b_j = 12$, $a_\nu = 1$ and $b_\nu = 10$), also assuming a large expected priors break size ($\gamma_j = 2.5$, $\delta_j = 500$, $\gamma_\nu = 1$ and $\delta_\nu = 250$).

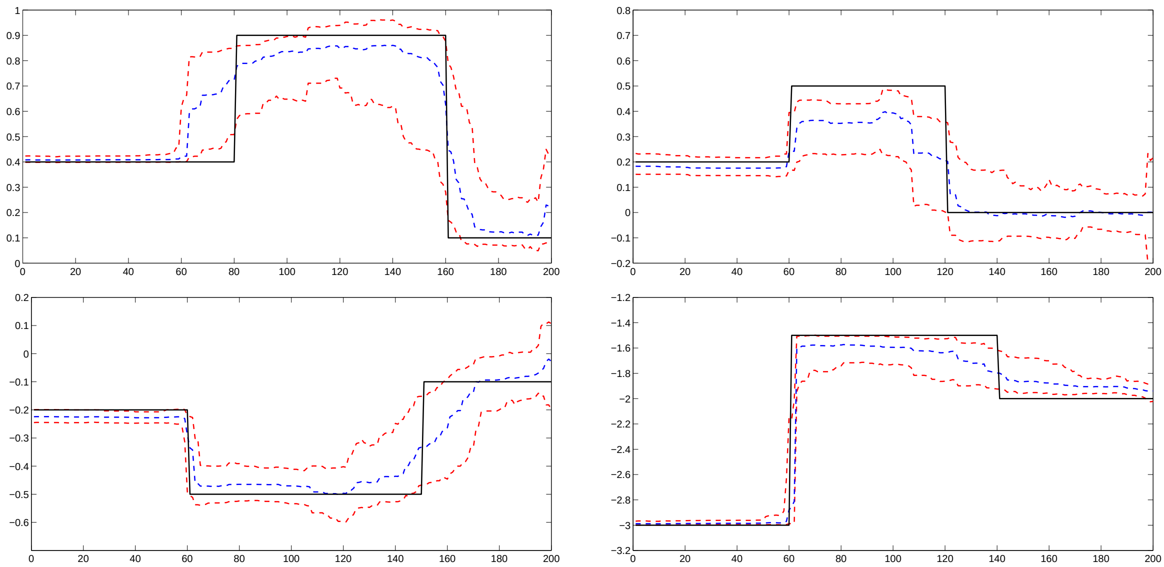
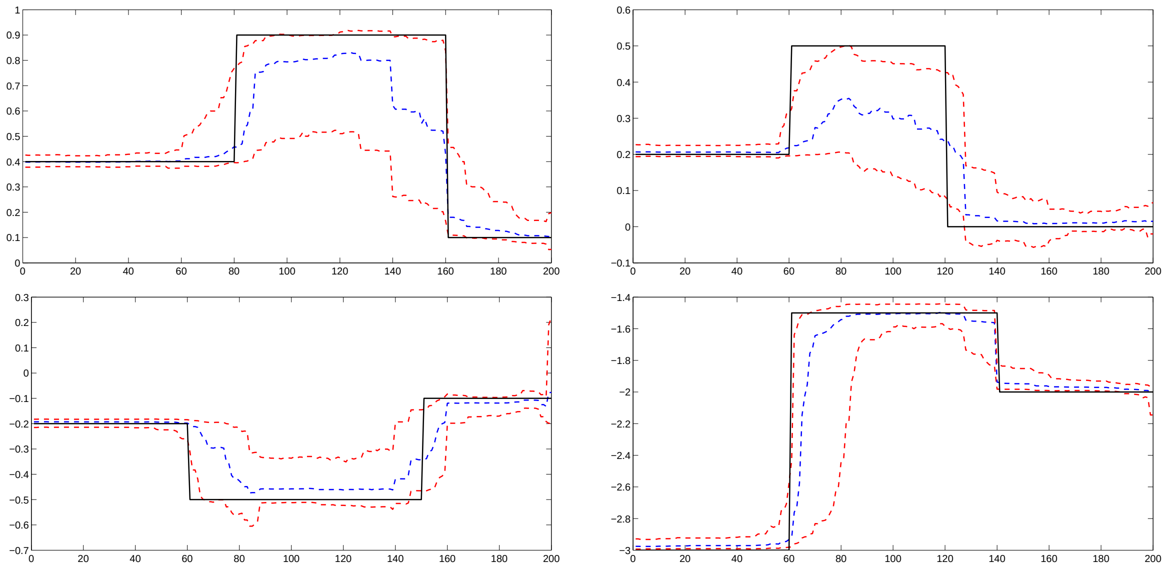


Figure C.4 shows the posterior estimates of $\beta_{i,t}$ for $i = 1, 2, 3$ and $\ln \sigma_t^2$ by assuming an higher prior probability of having a break and a smaller prior expected size of the breaks. The figures makes clear that much of the posterior medians deterioration showed in figure C.3 and C.2 would likely come from a higher expected size of the breaks. In particular assuming *a priori* small sized breaks will result in a less disperse posterior estimate and that the timing of breaks are mostly precisely estimated. A general pattern we observe is that when the prior settings correspond to a higher probability of smaller breaks than in the base case, the posterior estimates of $\beta_{1:3,t}$ and $\ln \sigma_t^2$ are consistent with those implied by the data generating process. However, by imposing ex-ante a larger size of breaks, the precision of posterior median estimates deteriorates. As a whole, the posterior estimates seems to be more sensible to prior hyper-parameters on the breaks size rather than to prior probabilities of breaks on itself.

Figure C.3: Posterior Estimates of the Time-Varying Parameters: Higher prior-break probabilities and large prior break size

This figure plots the posterior distributions of the parameters $\beta_{i,t}$ for $i = 1, 2, 3$ and $\ln \sigma_t^2$ together with the corresponding values implied by the DGP. The blue dashed line reports the median estimates of the parameters. The red dashed lines denote the 20th and 80th percentiles of the posterior distributions. The black solid line displays the values used to define the data generating process. Posterior results are now based on a larger prior-break probabilities for both betas and the (log of) idiosyncratic volatility ($a_j = 3.2$, $b_j = 6$, $a_\nu = 1$ and $b_\nu = 5$), also assuming a large expected priors break size ($\gamma_j = 2.5$, $\delta_j = 500$, $\gamma_\nu = 1$ and $\delta_\nu = 250$).




## C.2 Empirical Example

The second step of the prior sensitivity analysis is based on an empirical exercise. We base our results on the full B-TVB-SV model estimated on the original dataset of 23 stock and bond portfolios sorted on size, industry and maturity, and 9 macroeconomic risk factors (see Table 1 in the main text). The different prior specifications are those reported in Table C.3. Figure C.5 shows the distribution of posterior break probabilities for each of the macroeconomic risk factors, and averaged across the 23 stock and bond portfolios. The red dashed line corresponds to the posterior under the base case prior. The black line corresponds to the posterior under the "Large" case ($a_j = 3.2$, $b_j = 12$, $a_\nu = 1$ and $b_\nu = 10$), while the blue dot-dashed line represents the posterior distribution under the "Higher" case ($a_j = 3.2$, $b_j = 6$, $a_\nu = 1$ and $b_\nu = 5$). Figure C.5 makes clear that by assuming, *a priori* a higher probability of having a break does not lead to sensibly different posterior estimates of the instability of the betas in the full B-TVB-SV model. In fact, posterior estimates tend to largely overlap across explanatory macroeconomic factors. The same applies by looking at the posterior distribution of break probabilities for idiosyncratic variances. Figure C.6 reports the average posterior probabilities of having a break in $\ln \sigma_{i,t}^2$ under different priors. Again, posterior estimates tend to largely overlap under different priors.

As a further assessment we investigate the role of priors on break probabilities for $\ln \sigma_{i,t}^2$ in isolation. Figure C.7 shows the distribution of posterior break probabilities for each of the macroeconomic risk factors, and averaged across the 23 stock and bond portfolios. The prior structure for the betas is kept constant to the base case ($a_j = 3.2$, $b_j = 60$, and $\gamma_j = 0.5$, $\delta_j = 100$ for $j = 1, .., 10$). The red dashed line corresponds to the posterior under the base case prior ($a_\nu = 1$ and $b_\nu = 99$). The black line corresponds to the posterior under the

Figure C.4: Posterior Estimates of the Time-Varying Parameters: Higher prior-break probabilities and small prior break size

This figure plots the posterior distributions of the parameters $\beta_{i,t}$ for $i = 1, 2, 3$ and $\ln \sigma_t^2$ together with the corresponding values implied by the DGP. The blue dashed line reports the median estimates of the parameters. The red dashed lines denote the 20th and 80th percentiles of the posterior distributions. The black solid line displays the values used to define the data generating process. Posterior results are now based on a higher prior-break probabilities for both betas and the (log of) idiosyncratic volatility, while assuming a small expected priors break size.

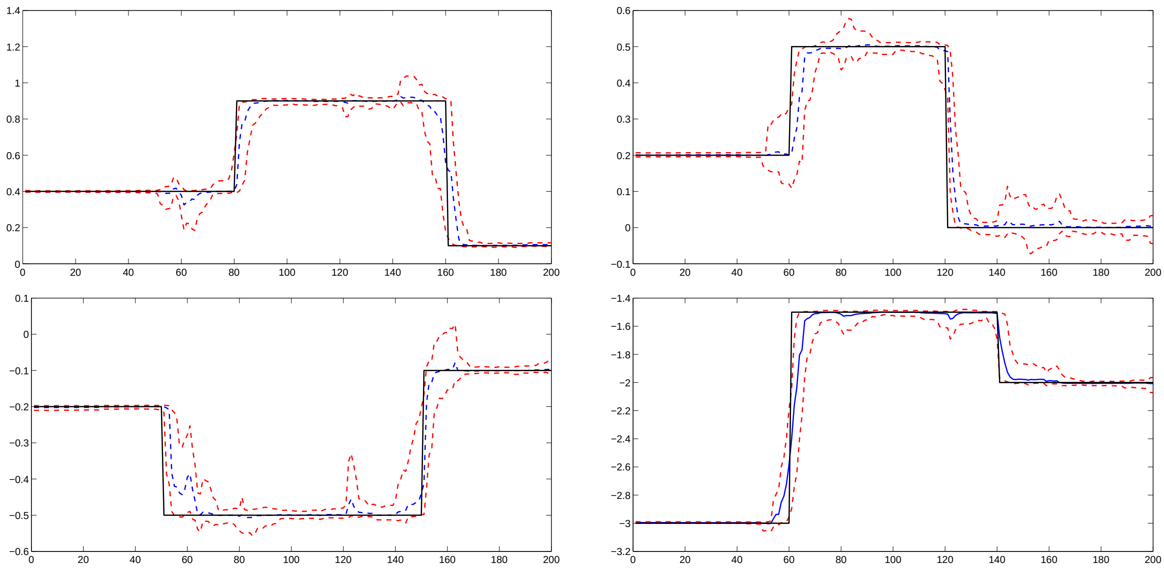


Figure C.5: Posterior Distributions of Break Probabilities of the Betas for Fixed Break Sizes

This figure plots the posterior distributions of the break probabilities for the betas averaged across portfolios and for each of the 9 macroeconomic factors depicted in Table B.1 in the main text. The red dashed line corresponds to the posterior under the base case prior. The black line corresponds to the posterior under the "Large" case ($a_j = 3.2$, $b_j = 12$, $a_\nu = 1$ and $b_\nu = 10$), while the blue dot-dashed line represents the posterior distribution under the "Higher" case ($a_j = 3.2$, $b_j = 6$, $a_\nu = 1$ and $b_\nu = 5$).

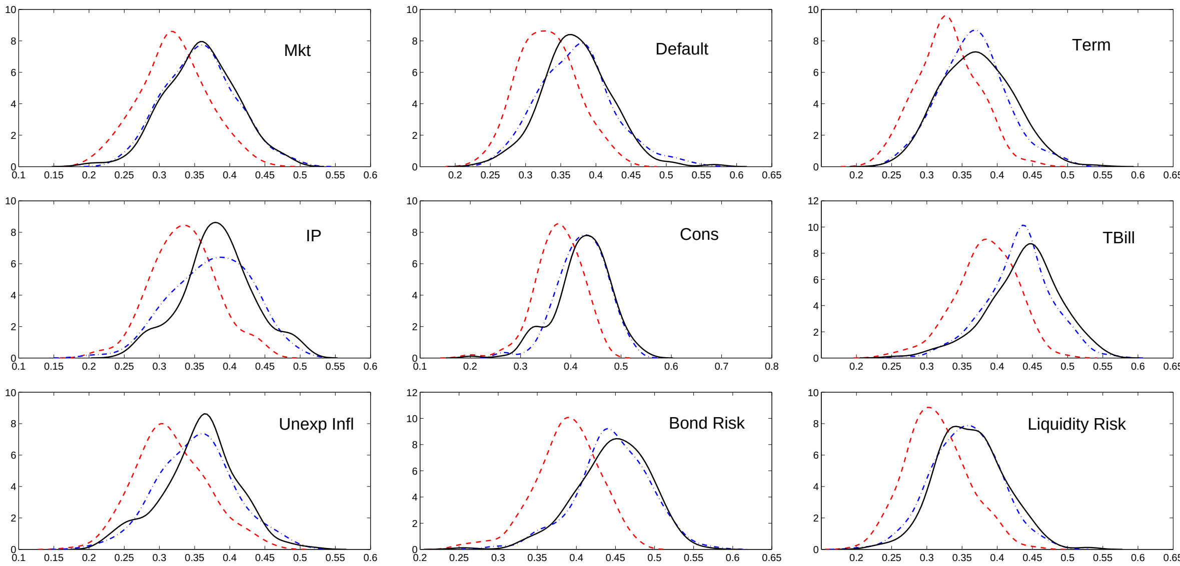
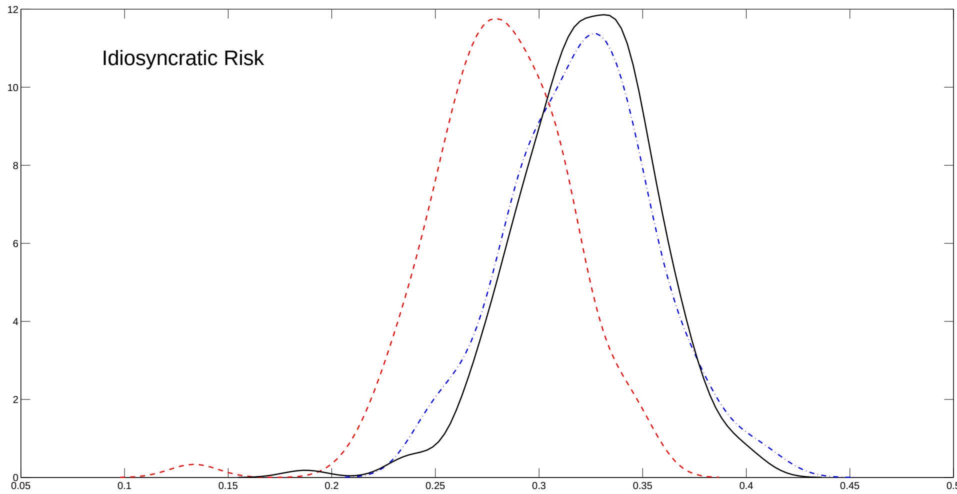


Figure C.6: Posterior Distributions of Break Probabilities of Idiosyncratic Volatility for Fixed Break Sizes

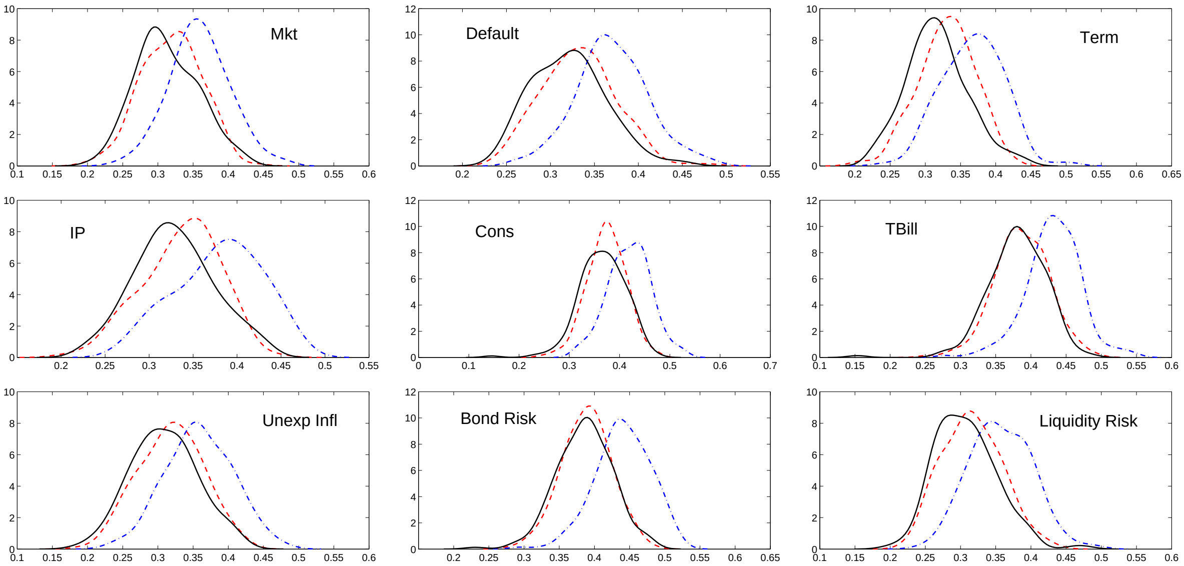
This figure plots the posterior distributions of the break probabilities for the (log of) idiosyncratic variances averaged across the 23 portfolios reported in Table 1 in the main text. The red dashed line corresponds to the posterior under the base case prior. The black line corresponds to the posterior under the "Large" case ($a_j = 3.2$, $b_j = 12$, $a_\nu = 1$ and $b_\nu = 10$), while the blue dot-dashed line represents the posterior distribution under the "Higher" case ($a_j = 3.2$, $b_j = 6$, $a_\nu = 1$ and $b_\nu = 5$).



"Large" case ($a_\nu = 1$ and $b_\nu = 10$), while the blue dot-dashed line represents the posterior distribution under the "Higher" case ($a_\nu = 1$ and $b_\nu = 5$). Figure C.7 makes clear that posterior estimates are rather robust with respect to different priors specifications for break probabilities on $\ln \sigma^2_{i,t}$. Interestingly, the data seem to be rather informative in defining the amount instability required by the dynamics of betas and idiosyncratic volatility. In fact, different priors do not lead to dramatically different results.

Figure C.7: Posterior Distributions of Break Probabilities of the Betas by Changing Priors on Idiosyncratic Volatility, and Keeping Fixed Break Sizes and Betas Prior Structure

This figure plots the posterior distributions of the break probabilities for the betas averaged across portfolios and for each of the 9 macroeconomic factors depicted in Table 1 in the main text. The prior structure on the betas is kept constant, while priors on $\ln \sigma^2_{i,t}$ change. The red dashed line corresponds to the posterior under the base case prior. The black line corresponds to the posterior under the "Large" case ($a_\nu = 1$ and $b_\nu = 10$), while the blue dot-dashed line represents the posterior distribution under the "Higher" case ($a_\nu = 1$ and $b_\nu = 5$).




# D In-Sample Posterior Inference on Stochastic Volatility

A number of studies in the finance literature have compared alternative models of time-varying volatility of asset returns (e.g. Hansen and Lunde 2005, Geweke and Amisano 2010, and Clark and Ravazzolo 2014). More recently Eisenstat and Strachan (2014) discuss estimation of volatility in the context of inflation, in particular whether it should be modelled as a stationary process or as a random walk. In our B-TVB-SV model when a break arrives, log-volatility follows a random walk. While such random walk assumption might be indeed useful for practical reasons, it can be criticized as inappropriate since it implies that the range of possible values of volatility is unbounded in probability in the limit, which is obviously something we do not observe in financial markets. On the other hand, stationary processes, say an AR(1) dynamics for log-volatility, are bounded in the limit, even though are close to be non-stationary at monthly frequencies, and also substantially increase the parameter space have to be estimated.

In this section, we report some in-sample properties of different specifications of the stochastic volatility component of our general model reported in Section 3 in the main text. The purpose is to investigate if the functional for volatility implied by the change-point dynamics sensibly affects the results in comparison of standard stationary and random walk specifications. For the model in-sample estimation we used the original Dataset of

23 portfolios of stocks sorted by size and industry, and bond portfolios sorted by maturity. Data are monthly and cover the sample period 1972:01 - 2011:12. The first ten years of data are used to calibrate the priors for each model.

For both discussion and estimation, we use a common observation equation specification for each of the excess returns $r_{t.t}$ on the 23 portfolios. For the sake of simplicity we assume a constant mean model. In general, we write this as

$$r_{i,t} = \mu + \sigma_{it}\epsilon_{i,t} \qquad i = 1, ..., N,$$

We restrict the discussion to two specifications, namely our change-point dynamics as in Section 3, and a standard AR(1) stationary process. The change-point specification for log-volatility, is defined as (see the main text for more details)

$$\ln(\sigma^2_{i,t}) = \ln(\sigma^2_{i,t-1}) + \kappa_{iv,t}\upsilon_{i,t} \quad \text{with} \quad \upsilon_{i,t} \sim N(0, q_i) \quad \text{and} \quad \kappa_{iv,t} = \begin{cases} 1 & \pi_{iv} \\ 0 & 1 - \pi_{iv} \end{cases} \tag{D.13}$$

The alternative stationary specification we consider is a standard AR(1) dynamics

$$\ln(\sigma^2_{i,t}) = (1 - \delta_i)\,\overline{\ln(\sigma^2_i)} + \delta_i \ln(\sigma^2_{i,t-1}) + \upsilon_{i,t} \qquad \text{with} \qquad \upsilon_{i,t} \sim N(0, q_i) \tag{D.14}$$

with $\overline{\ln(\sigma^2_i)}$ the long-run mean, and $\delta_i$ the asset specific persistence parameter of log-volatility. Clearly both the change-point model and the AR(1) nests a random walks dynamics when $\pi_{iv} = 1$ and $\delta_i = 1$, respectively. Figure 6 shows the median estimates of $\ln(\sigma^2_{i,t})$ according to (D.13) and (D.14) respectively. We report the results for the size-sorted portfolios for the sake of readability. The blue line corresponds to the stationary AR(1), while the red line is the median estimate under the B-TVB-SV model. This figure makes clear that both of the models specifications helps to capture spikes in conditional volatility, for instance around the period 2000/2002, across different assets. In other words, at least in finite samples, there is not clear benefit in using a stationary dynamics as the AR(1) as opposed to the full B-TVB-SV model. One potential reason is that, indeed, at the monthly frequency, the AR(1) dynamics of log-volatility is close to be non-stationary, i.e. $\delta_i = 1$. Figure **??** shows the posterior distribution of the persistence parameters $\delta_i$ across the same set of size-sorted portfolios reported in Figure 6. Indeed, shocks to the AR(1) log-volatility turn out to have a largely persistent effect. The average posterior median estimate of $\delta_i$ is well above 0.9 on a monthly frequency.

A general pattern we observe is that, at least in finite samples, both a highly persistent AR(1) and a change-point dynamics may help to capture the same dynamic features of log-volatility. The question is now why we argue the latter might be indeed better to fit the data. We compared the log-marginal likelihoods of the specifications (D.13)-(D.14) together with a standard random walk dynamics. Table D.4 shows the results. The full B-TVB-SV model delivers the highest log-marginal likelihood across all the size-sorted portfolios. Interestingly, the data are clearly in favor of the stationary AR(1) dynamics as opposed to the random walk one.

As a whole, posterior estimates of log-volatility would not radically change by using a more standard stationary AR(1) dynamics. The latter however, is strongly rejected by the data in typical set of 40 years of post-war data on size-sorted stocks. Interestingly, the full B-TVB-SV closely behaves as a highly persistent, stationary, AR(1) process (of course, this is true in finite samples and not asymptotically).

# E Variance Decomposition Tests

We use the posterior densities of the time series of factor loadings and risk premia to perform a number of tests that allow us to assess whether a posited asset pricing framework may explain an adequate percentage of excess asset returns. (5) decomposes excess asset returns in a component related to risk, represented by the term $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t-1}$ plus a residual $\lambda_{0,t} + e_{i,t}$. In principle, a multi-factor model is as good as the implied percentage of total variation in excess returns explained by its first component, $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t-1}$. However, here we should recall that even though (5) refers to excess returns, it remains a statistical implementation of the framework in (4). This implies that in practice it may be naive to expect that $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t-1}$ be able to explain much of the

Figure D.8: Posterior Median Estimates of Log-Volatility Under the Change-Point Dynamics vs. Stationary AR(1)

This figure plots the posterior median estimates of the log-volatility for a set of size-sorted portfolios across the period 1972:01 - 2011:01. The blue line corresponds to the stationary AR(1), while the red line is the median estimate under the B-TVB-SV model. Prior hyper-parameters are trained in both cases by using a pre-sample period of ten years.

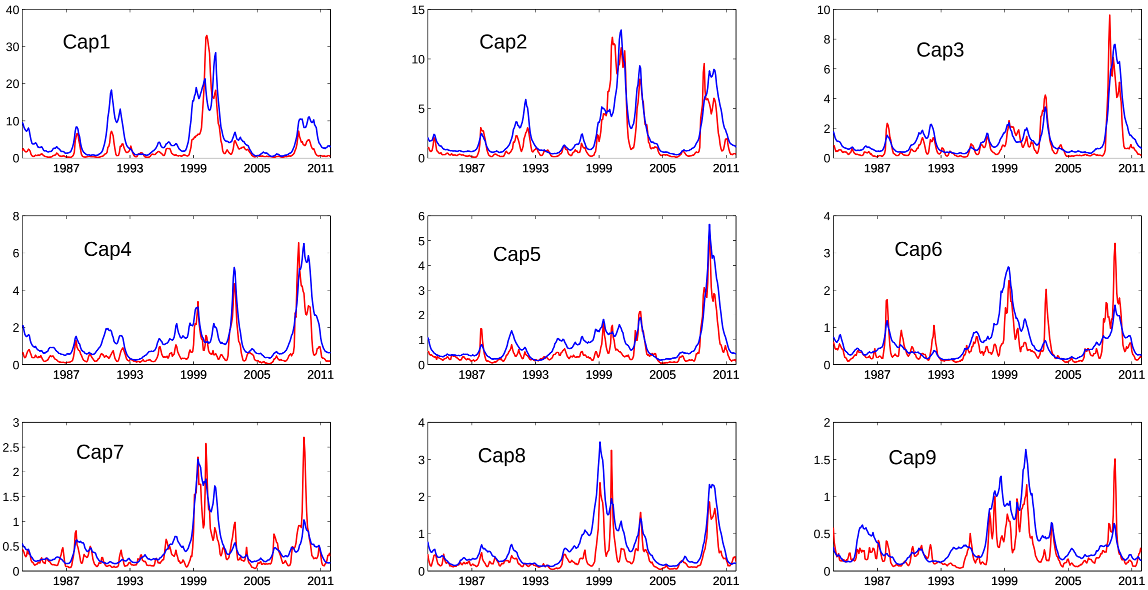


Figure D.9: Posterior Estimates of the Persistence Parameters for the AR(1) Log-Volatility

This figure plots the posterior distribution estimates of the persistence parameters $\delta_i$ for the log-volatility for a set of size-sorted portfolios across the period 1972:01 - 2011:01. The blue line corresponds to the stationary AR(1), while the red line is the median estimate under the B-TVB-SV model. Prior hyper-parameters are trained in both cases by using a pre-sample period of ten years.

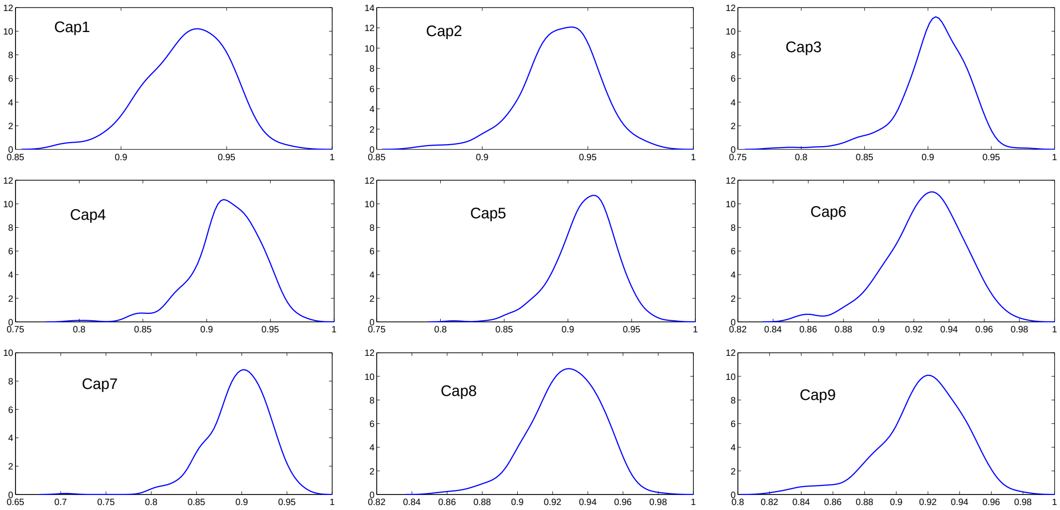



Table D.4: Log-Marginal Likelihoods Across Alternative Stochastic Volatility Specifications

This table the values of the log-marginal likelihoods for different specifications of stochastic volatility. The values of log marginal likelihoods are reported for ten stocks portfolios sorted on size. *Change-Point* stands for the full model proposed in the main text, while *Stationary* and *Random Walk*, respectively represents a model with a stationary and random walk dynamics for the stochastic volatility process.

| | **10 Size-Sorted Portfolios, Value Weighted** | | |
|---|---|---|---|
| | **Change-Point** | **Stationary** | **Random Walk** |
| Decile 1 | -534.140 | -613.760 | -805.997 |
| Decile 2 | -444.477 | -598.132 | -776.736 |
| Decile 3 | -307.812 | -479.321 | -734.209 |
| Decile 4 | -279.771 | -460.944 | -733.441 |
| Decile 5 | -232.713 | -415.034 | -717.713 |
| Decile 6 | -217.594 | -432.015 | -719.466 |
| Decile 7 | -168.411 | -445.438 | -705.954 |
| Decile 8 | -148.350 | -312.585 | -704.538 |
| Decile 9 | -96.393 | -433.774 | -691.064 |
| Decile 10 | -43.428 | -222.075 | -683.426 |

variability in excess returns. A more sensible goal seems to be that $\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t-1}$ ought to at least explain the *predictable* variation in excess returns. We therefore follow earlier literature, such as Karolyi and Sanders (1998), and adopt the following approach. First, the excess return on each asset is regressed onto a set of $M$ instrumental variables that proxy for available information at time $t-1$, $\mathbf{Z}_{t-1}$,

$$x_{i,t} = \theta_{i0} + \sum_{m=1}^{M}\theta_{im}Z_{m,t-1} + \xi_{i,t}, \tag{E.15}$$

to compute the sample variance of fitted values,

$$Var[P(x_{it}|\mathbf{Z}_{t-1})] \equiv Var\left[\hat{\theta}_{i0} + \sum_{m=1}^{M}\hat{\theta}_{im}Z_{m,t-1}\right], \tag{E.16}$$

where the notation $P(x_{it}|\mathbf{Z}_{t-1})$ means "linear projection" of $x_{it}$ on a set of instruments, $\mathbf{Z}_{t-1}$. Second, for each asset $i = 1, ..., N$, a time series of fitted (posterior) risk compensations, $\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t-1}$, is regressed onto the instrumental variables,

$$\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t-1} = \theta'_{i0} + \sum_{m=1}^{M}\theta'_{im}Z_{m,t-1} + \xi'_{i,t} \tag{E.17}$$

to compute the sample variance of fitted risk compensations:

$$Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1}\right)\right] \equiv Var\left[\hat{\theta}'_{i0} + \sum_{m=1}^{M}\hat{\theta}'_{im}Z_{m,t-1}\right]. \tag{E.18}$$

The predictable component of excess returns in (E.15) not captured by the model is then the sample variance of the fitted values from the regression of the residuals $\hat{\xi}_{i,t}$ on the instruments:

$$Var\left[\hat{\xi}_{i,t}\right] = Var\left[P\left(\lambda_{0,t} + e_{i,t}|\mathbf{Z}_{t-1}\right)\right]. \tag{E.19}$$

At this point, it is informative to compute and report two variance ratios, commonly called $VR1$ and $VR2$, after Ferson and Harvey (1991):

$$VR1 \equiv \frac{Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var[P(x_{it}|\mathbf{Z}_{t-1})]} > 0 \tag{E.20}$$

$$VR2 \equiv \frac{Var\left[P\left(\lambda_{0,t} + e_{i,t}|\mathbf{Z}_{t-1}\right)\right]}{Var[P(x_{it}|\mathbf{Z}_{t-1})]} > 0. \tag{E.21}$$

VR1 should be equal to 1 if the multi-factor model is correctly specified, which means that all the predictable variation in excess returns is captured by variation in risk compensations; at the same time, VR2 should be equal to zero if the multi-factor model is correctly specified. Importantly, when these decomposition tests are implemented using the estimation outputs obtained from our B-TVB-SV framework, drawing from the joint posterior densities of the factor loadings $\beta_{ij,t-1}$ and the implied risk premia $\lambda_{j,t}$, $i = 1, ..., N$, $j = 1, ..., K$, and $t = 1, ..., T$, and holding the instruments fixed over time, it is possible to compute VR1 and VR2 in correspondence to each of such draws and hence obtain their posterior distributions.[1]

Finally, the predictable variation of returns due to the multi-factor model may be further decomposed into the components imputed to each of the individual systematic risk factors, by computing the factoring of $Var[P(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1})]$ as

$$\sum_{j=1}^{K}Var\left[P\left(\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1}\right)\right] + \sum_{j=1}^{K}\sum_{k=1}^{K}Cov[P\left(\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1}\right), P\left(\lambda_{k,t}\beta_{ik,t-1}|\mathbf{Z}_{t-1}\right)] \tag{E.22}$$

[1]Notice that $VR1 = 1$ does not imply that $VR2 = 0$ and viceversa, because

$$Var[P(x_{it}|\mathbf{Z}_{t-1})] \neq Var\left[P\left(\sum_{j=1}^{K}\hat{\lambda}_{j,t}\hat{\beta}_{ij,t-1}|\mathbf{Z}_{t-1}\right)\right] + Var\left[P\left(r_{i,t} - \hat{\theta}_{i0} - \sum_{m=1}^{M}\hat{\theta}_{im}Z_{m,t-1}|\mathbf{Z}_{t-1}\right)\right].$$

and tabulating $Var\left[P\left(\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1}\right)\right]$ for $j=1,...,K$ as well as the residual factor $\sum_{j=1}^{K}\sum_{k=1}^{K}Cov[\,P\left(\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1}\right),$ $P\left(\lambda_{k,t}\beta_{ik,t-1}|\mathbf{Z}_{t-1}\right)]$ to pick up any interaction terms. Note that because of the existence of the latter term, the equality

$$\sum_{j=1}^{K}\frac{Var\left[P\left(\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t-1}|\mathbf{Z}_{t-1}\right)\right]}=1 \tag{E.23}$$

fails to hold, i.e., the sum of the $K$ risk compensations should not equal the total predictable variation from the asset pricing model because of the covariance among individual risk compensations. This derives from the fact that even though in (1) the risk factors are assumed to be orthogonal, this does not imply that their time-varying total risk compensations ($\lambda_{j,t}\beta_{ij,t-1}$ for $j=1,...,K$) should be orthogonal.