A Bayesian Multi-Factor Model of Instability in Prices and Quantities of Risk in U.S. Financial Markets

Authors	Massimo Guidolin, Francesco Ravazzolo, and Andrea Donato Tortora
Working Paper Number	2011-003A
Creation Date	January 2011
Citable Link	https://doi.org/10.20955/wp.2011.003
Suggested Citation	Guidolin, M., Ravazzolo, F., Tortora, A.D., 2011; A Bayesian Multi-Factor Model of Instability in Prices and Quantities of Risk in U.S. Financial Markets, Federal Reserve Bank of St. Louis Working Paper 2011-003. URL https://doi.org/10.20955/wp.2011.003

A Bayesian Multi-Factor Model of Instability in Prices and Quantities of Risk in U.S. Financial Markets*

Massimo Guidolin†, Francesco Ravazzolo‡ and Andrea Donato Tortora§

This version: January 2011

Abstract

This paper analyzes the empirical performance of two alternative ways in which multi-factor models with time-varying risk exposures and premia may be estimated. The first method echoes the seminal two-pass approach advocated by Fama and MacBeth (1973). The second approach extends previous work by Ouysse and Kohn (2010) and is based on a Bayesian approach to modelling the latent process followed by risk exposures and idiosynchratic volatility. Our application to monthly, 1979-2008 U.S. data for stock, bond, and publicly traded real estate returns shows that the classical, two-stage approach that relies on a nonparametric, rolling window modelling of time-varying betas yields results that are unreasonable. There is evidence that all the portfolios of stocks, bonds, and REITs have been grossly over-priced. On the contrary, the Bayesian approach yields sensible results as most portfolios do not appear to have been misspriced and a few risk premia are precisely estimated with a plausibile sign. Real consumption growth risk turns out to be the only factor that is persistently priced throughout the sample.

Key words: Bayesian estimation, Latent jumps, Stochastic volatility, Linear factor models.

JEL codes: G11, C53.

1. Introduction

What are the key features of the general pricing mechanisms—sometimes called the pricing kernel—underlying the observed prices for securities that can be traded in the capital market? This is of course one of the key questions in financial economics and has originated an entire field devoted to this very question, asset pricing. However, a related but not less important question is: What are the most appropriate methods available to researchers as well as practitioners to learn about the pricing kernel that underlies the observed cross-section of asset prices and returns? Our paper offers a contribution to the voluminous

*We are grateful to seminar participants at the Norges Bank, the IREBS 2009 Conference on Real Estate Economics and Finance, the JEuBES Japanese-European Bayesian Econometrics and Statistics Meeting (2009) and the 4th International Conference on Computational and Financial Econometrics in London (CFE'10) for helpful comments and suggestions. Massimo Guidolin acknowledges financial support from the Center for Analysis of Investment Risk at the University of Manchester. Andrea Donato Tortora acknowledges financial support from the Marie Curie Early Stage Training Programme. The views expressed in this paper are our own and do not necessarily reflect those of Norges Bank.

†Federal Reserve Bank of St. Louis and Bocconi University, IGIER. E-mail: massimo.guidolin@unibocconi.it.

‡Norges Bank, Research Department. E-mail: Francesco.Ravazzolo@Norges-Bank.no.

§Bocconi University, Milan. E-mail: andrea.tortora@phd.unibocconi.it.

literature that has tackled both questions specializing to a particular set of models (of the pricing kernel), of empirical methods, and applying novel methods to an interesting application.[1] With reference to an application to 25 years of monthly data on excess returns on 28 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 approach that allows for instability in factor exposures and risk premia delivers encouraging (at times, surprising) and sensible results.

The paper has three building blocks. First, our paper postulates and estimates using two alternative empirical approaches a standard multifactor asset pricing model (MFAPM) in which the proposed risk factors consists of shocks to observable macroeconomic variables that appear to be commonly tracked by researchers, policy-makers, and the press (e.g., aggregate market returns, the rate of growth of industrial production, changes in the unemployment rate, the spread between long- and short-term nominal rates, etc.). Going back to the seminal paper by Chen, Roll and Ross (1986) there is of course an ever expanding literature that has worked with such a class of models, to test their performance and to ask applied questions. In particular, Ferson and Harvey (1991) extended the early work on MFAPMs to incorporate the case of time-varying risk premia and exposures. In general terms, a MFAPM has a very simple structure: the risk premium on any asset or portfolio is decomposed as the sum of a certain number (K) of products between risk exposures (also called betas) to each of the factors and the associated unit price of the factor. The difference at each point in time between actual, realized excess returns and the risk premium implied by the model is called residual or idiosyncratic risk and it is supposed to pick up all the variation in excess returns that is specific to each individual portfolio.

Second, our paper jointly uses data on publicly traded stock, bond, and real estate securities (or traded funds invested in the underlying securities), instead of focussing on only one of these asset classes. Therefore, our paper relates to a vast literature that has examined the empirical performance of MFAPMs across asset classes. For instance, Chan, Hendershott and Sanders (1990) have shown that MFAPMs that include predetermined macroeconomic factors explain a significant proportion of the variation in equity real estate investment trusts (henceforth REITs) returns. Karolyi and Sanders (1998) have extended this evidence and allowed for time-varying risk premia and betas in a similar framework to the seminal paper by Ferson and Harvey (1991).

Third, our key contribution consists of comparing the heterogeneous results derived from two alternative "implementations" of a rather standard, Ferson and Harvey (1991)-style MFAPM.[2] The first approach follows the now standard, Fama and MacBeth (1973) two-stage methodology, first proposed for the plain vanilla CAPM but then extended to the wider class of linear factor models. Basically, Fama-MacBeth's approach uses a first set of rolling window, time series regressions to obtain least-square estimates of the risk exposures, followed by a second-pass set of cross-sectional (across assets or portfolios) regressions

[1] A general introduction to linear factor models, their relationship to the general notion of the pricing kernel in asset pricing, and related estimation methods can be found in textbooks such as Cochrane (2005), and Singleton (2006).

[2] The use of the concept of implementation intends to stress that in a given asset pricing model—for instance, in which risk exposures and premia are in principle time-varying—alternative estimation methodologies can be deployed to learn from the data the dynamics over time of betas and risk premia.

that—using the first-pass rolling window betas as inputs—derives time-varying (monthly) estimates of the associated risk premia. The problems with this methodology are notorious and well-visible to any applied econometrician: 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 clearly clashes with them being least squares estimates (and therefore random sample statistics). Obviously, unless additional (and generally dubious) assumptions are introduced, this creates a potentially enormous problem with generated regressors being used in the second-step, which tends to make invalid most the inferential statements commonly made when the resulting error-in-variables problems are ignored. Fama-MacBeth's approach also suffers from another problem: although now common, identifying time-variation in risk exposures and risk premia with a need to perform rolling window least square estimation is surely robust (because, in a way, nonparametric) but always arbitrary and often unsatisfactory. As a result of these drawbacks of Fama-MacBeth's approach, we follow a different path based on two steps:

- Time variation in risk exposures and premia is explicitly modelled as a break-point process; the parameters of interest (βs and log-idiosyncratic volatilities, see Section 2 for details and definitions) are constant unless a break-point variable (κ) 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; the break-point variable κ takes a value of one, signalling the occurrence of a jump, with some probability (π) which is itself an estimable parameter; finally, the breaks themselves are latent, which implies that the data ought to be used also to make inferences on the dates and magnitudes of the breaks.

- The model is estimated using a Bayesian approach that not only is numerically practical, and as usual allows a researcher to feed her own priors on the quantities of interest in the estimation problem, but also allows us to naturally overcome the issues with generator regressors described above.[3]

We model both factor sensitivities and idiosyncratic volatility as latent stochastic processes within a Bayesian framework by means of the mixture innovation approach as in Giordani, Kohn and van Dijk (2007) and Giordani and Kohn (2008). Furthermore, we estimate the sequence of risk premia following Ouysse and Kohn (2010) to consistently overcome the problems with generated regressors. We show that this approach helps reduce the extent of variations in estimated risk premia. Notice that our claim here is not that our Bayesian, latent variables approach is the only possible solution to the issues plaguing the classical Fama-MacBeth's approach.[4] However, our paper shows—for a case in which Fama-MacBeth's method can be illustrated to fail and to lead to rather nonsensical economic results—that the payoff may be substantial and that the extra computational issues to be dealt with are certainly worth it.

Our main results can be summarized as follows. The general lesson is that—at least in our application and with our data—a standard two-stage Fama-MacBeth approach yields unreasonable economic

[3]Although this is hard to prove, there are considerable doubts on whether and how such a break-point MFAPM model could be estimated using classical methods, i.e., relying on the likelihood of the data only.

[4]For instance, Cochrane (2005) features a textbook-level discussion of how the Generalized Method of Moments (GMM) may be deployed to overcome the effects of the errors-in-variables problem within typical Fama-MacBeth applications.

implications and is drastically rejected in a statistical sense. First, our Bayesian estimates of the loadings (the betas) are considerably smoother than the rolling window, Fama-MacBeth ones, which are instead subject to massive instability, often impossible to interpret. This is a seemingly counter-intuitive result: even though a Bayesian model with latent breaks formally allows the risk exposures to be subject to jumps over time, the resulting posterior densities are actually smoother than what one could retrieve using a naïve rolling window estimation procedure. Second, the two-step Fama-MacBeth case leads to the rather implausible finding that all the 28 test portfolios display large and negative estimates of a parameter (β_0) that measures the existence of abnormal returns that cannot be justified by exposure to systematic risks. This means that all of our portfolios would have been systematically and persistently over-priced during our sample period. This would be an overwhelming indication of irrational exuberance of the U.S. market over a 25 year long sample. On the contrary, in the Bayesian case, the values of the posterior means of the β_0 parameters as well as their signs are economically plausible. Even more interestingly, once we take into account their 90% confidence bands, it is clear that in 28 out of 28 portfolios, the reported posterior means fail to be precisely estimated, in the sense that the 90% band always includes a zero abnormal return during the period 1983-2008. However, there is some evidence that durable stocks (especially after 2000), high-tech stocks (throughout the entire sample, a persistent tech bubble), small capitalization stocks, and mortgage REITs (especially after 2000) have gone through episodes of excess returns substantially below the levels justified by their risk exposures and therefore can be considered to have been over-priced.

Third, the Fama-MacBeth approach shows that idiosyncratic risk is large for most portfolios investigated, a sign that the two-pass method provides a very rough explanation to the 28 time series of excess returns data. In the Bayesian case, when all the uncertainty is taken into account, there is no strong evidence of trends in idiosyncratic risk, even though the plots for the individual portfolios show some evidence of a peak in the early 2000s and some sign of a growing trend towards the end of our sample, consistently with some earlier literature (see e.g., Campbell, Lettau, Malkiel and Xu, 2001, and Zhang, 2010).

Fourth, under Fama-MacBeth, none of the risk premia turns out to be precisely estimated, apart from the risk premium on a riskless yield curve slope factor. However, this premium is significantly negative, which—assuming that a steeper yield curve forecasts an improvement in business cycle conditions —is puzzling. The only quantity for which there is compelling evidence is λ_0, which is a measure of abnormal returns not justified by risk exposures. λ_0 is estimated to be negative and massive (-3.92% per month), with an essentially zero p-value. In a sense, all that a standard Fama-MacBeth approach reveals is that U.S. asset data contain strong evidence of structural misspricings in the form of negative abnormal returns. On the opposite the Bayesian (posterior median) estimates of the risk premia are considerably more stable and—more importantly—a few of them are precisely estimated. Here one result is striking: with reference to the full sample, real consumption growth yields the only precisely estimated risk premium and it displays the expected positive sign. Also, the Bayesian design gives evidence of moderate but precisely estimated misspricings, with an average posterior median of λ_0 equal to 0.37% per month. However, these misspricings now correspond to over-pricing of the risk factors captured by the model and to an (indeed modest) structural under-pricing of the assets over time. Moreover, a sub-sample (1983-1992) is found over which the average posterior median of λ_0 is small (0.26%) and fails to be significant.

A closely related paper that deserves discussion is Ouysse and Kohn (2010) who use Bayesian variable selection and Bayesian model averaging in inferring factors for MFAPMs. Modelling these sources of uncertainties increases estimation accuracy, particularly in finite samples, as it is often the case. Ouysse and Kohn find new empirical evidence of time-varying risk premiums with higher and more volatile expected compensation for bearing systematic risk during contraction phases. However, Ouysse and Kohn limit themselves to the analysis of an unconditional APT model and focus only on stock portfolios, excluding bond and real estate data. Moreover, they do not investigate the issue of parameter instability, which is instead the main goal of our paper.

The remainder of the paper is organized as follows. Section 2 outlines our theoretical MFAPM and describes both the classical Fama-MacBeth implementation based on rolling window estimates of risk exposures and premia, and our Bayesian approach to estimate models with latent stochastic breaks and variances. The Section also briefly presents a few standard (variance) ratios used in the literature to evaluate the "economic" fit of MFAPMs. Section 3 describes the data used in the empirical analysis. Section 4 reports the main empirical results and performs a comparison between two-pass Fama-MacBeth results and Bayesian posterior results. Section 5 performs robustness checks, concerning the MFAPM implementation as well as the priors, as customary. The concluding section summarizes our findings.

2. **Research Design and Methodology**

2.1. *The Asset Pricing Framework*

Our research design is based on an extension of the multi-factor models first introduced by Ferson and Harvey (1991) in the asset pricing literature and applied to a variety of contexts and empirical research questions (see e.g., Chan, Henderhott and Sanders, 1990, and Karolyi and Sanders, 1998, for an application to public real estate investment vehicles; Ferson, 1996, for an application to mutual fund performance evaluation; Ilmanen, 1995, for an application to bond pricing, etc.). The common thread of these papers is their attempt at connecting return predictability to variation in the expected compensation for risk. In particular, the literature on multifactor asset pricing models (MFAPMs) posits a linear relationship between asset returns and a set of macroeconomic factors that are assumed to capture business cycle effects on beliefs and/or preferences (as summarized by a pricing kernel factor with time-varying properties) and hence on risk premia. These macroeconomic factors with general effects on the kernel are typically market portfolio (i.e., aggregate wealth) returns, the default spread on corporate bond yields, the term spread incorporated in the riskless (Treasury) yield curve, and changes in the rate of growth of industrial production (see Chen, Roll and Ross, 1986, and Liu and Mei, 1992, with specific reference to real estate finance applications). If we call the process of the (shocks to) macroeconomic risk factors $F_{j,t}$ $(j = 1, ..., K)$ and $r_{i,t}$ the period *excess* return on asset or portfolio $i = 1, ..., N$, then a typical MFAPM can be written as:

$$r_{i,t} = \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \epsilon_{i,t}, \tag{1}$$

where it is customary to assume that $E[\epsilon_{i,t}] = E[F_{j,t}] = E[\epsilon_{i,t}F_{j,t}] = 0$ for all $i = 1, ..., N$ and $j = 1, ..., K$. The $r_{i,t}$ are returns in excess of the risk-free rate proxied by the 1-month T-bill. Normally, the advantage of MFAPMs such as (1) consists of the fact that a number of systematic factors $K << N$ may prove useful to efficiently capture relatively large portions of the variability in asset returns. Importantly, even though the notation $\beta_{ij,t}$ lets us understand that the factor loadings are allowed to be time-varying, such patterns of time variation are in general left unspecified.

In the conditional version of Ross' (1976) APT or Merton's (1973) ICAPM, the expected excess return on asset i over the interval $[t-1, t]$ (the risk premium on asset i) may then be related to its "betas" (i.e., factor loadings measuring the exposure of asset i to each of the priced, systematic risk factors) and the associated unit risk premia (i.e., average return compensations for unit exposure to risk):

$$E[r_{i,t}|\mathbf{Z}_{t-1}] = \lambda_0(\mathbf{Z}_{t-1}) + \sum_{j=1}^{K} \beta_{ij,t}\lambda_j(\mathbf{Z}_{t-1}), \tag{2}$$

where both the betas and the risk premia are conditional on the information publicly available at time t, here summarized by the $M \times 1$ vector of "instruments" $\mathbf{Z}_t$.

2.2. *Standard Fama-MacBeth Implementation*

The framework in (1)-(2) just describes a general conditional pricing framework that is actually known to hold under a variety of assumptions and conditions.[5] However, a variety of alternative methodologies have been proposed to perform three related tasks which impinge on the empirical performance of (1)-(2): (i) how many factors ought to be selected, i.e., picking an appropriate K; (ii) given K, devising a methodology to be able to rank competing factors and selecting only those that are "required" by the data; (iii) estimating the factor loadings $\{\beta_{ij,t}\}$ (over time and for each possible pair i, j) and the risk premia λ_{jt} (over time and for each possible j). These tasks are logically distinct from the formulation of the asset pricing framework and—albeit their optimal implementation affects our ability to learn about the asset pricing questions under investigation—they have an exquisite statistical nature. In this paper we align ourselves to a number of related papers in the empirical finance literature (see e.g., Chen, Roll and Ross, 1986) as far (i)-(ii) are concerned—which means that we pre-select both K and which specific macroeconomic risk factors ought to be considered in the light of the existing literature—and pursue two alternative econometric approaches with regard to task (iii).[6]

The first approach is the classical, two-stage procedure *à* la Fama and MacBeth (1973) also used by Ferson and Harvey (1991).[7] In the first stage, for each of the assets, the factor betas are estimated using

[5]For instance, standard arguments in Cochrane (2005) show that (1)- (2) holds when the stochastic discount factor can be written as an exact linear function of the systematic risk factors F_1, F_2, ..., F_K.

[6]See Groen, Paap and Ravazzolo (2009) for an example on how to extend the time-varying framework in (1) to incorporate the variable selection and model averaging methods proposed by Ouysse and Kohn (2010). More efficient algorithms than in Groen, Paap and Ravazzolo (2009) for variable search as in Hofmann, Gatu, Kontoghiorghes (2007), and Kapetanios (2007) could also find application.

[7]In fairness we need to add that Ferson and Harvey (1991) have explored a range of alternative beta estimation techniques, including conditional betas estimated from regressions on past information variables, sixty-month rolling betas regressed on

time-series regressions from historical excess returns on the assets and economic factors. That is, for month t, we estimate equation (1) using the previous sixty months (ranging from $t-61$ to $t-1$) in order to obtain estimates for the betas, $\hat{\beta}^{60}_{ij,t-1}$. This time-series regression is updated each month. In the second stage, we estimate a cross-sectional regression, for each month, using ex-post realized excess returns

$$r_{i,t} = \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t} \hat{\beta}^{60}_{ij,t-1} + \zeta_{i,t} \qquad i = 1, ..., N, \tag{3}$$

(and for each $t = 60, ..., T$). In (3) $\lambda_{0,t}$ is the zero-beta (abnormal) excess return and the $\lambda_{j,t}$s are proxies for the factor risk premiums on each month, $j = 1, ..., K$. Notice that $\lambda_{0,t}$ should equal zero $\forall t$ if the model is correctly specified, because in the absence of arbitrage all zero-beta assets should command a rate of return that equals the risk-free rate. Tests of multi-factor models evaluate the importance of these economic risk variables by evaluating whether their risk premiums are priced or whether, on average, the (second-stage, estimated) coefficients $\hat{\lambda}_{j,t}$ are significantly different from zero.

2.3. *A New Bayesian Estimation Approach*

Although widely used in the applied finance literature, the classical two-stage Fama-MacBeth approach has a number of obvious statistical drawbacks. To name only two, first—it is clear that the second stage multivariate regression in (3) 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 $\hat{\beta}^{60}_{ij,t-1}$ covary 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 seems to contain a general presumption that during business cycle downturns both the quantity of risk (here captured by the size of the betas) and the unit risk prices tend to increase, simply because recessions are characterized by higher systematic uncertainty as well as by lower "risk appetite" (for instance, in a Campbell and Cochrane's, 1999, habit-formation asset pricing framework). Second, the need to perform the necessary estimation to implement (1)-(2) in two distinct stages that use rolling windows to capture parameter instability is not only *ad hoc* but also largely inefficient because the lack of more specific parametric forms makes it 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.

Clearly, both issues are tackled by any full-information estimation method that would avoid using estimates of the first-stage betas as if these were observed variables constant in repeated samples and that would take into account the existence of time-varying factor loadings and idiosyncratic variance in specific parametric forms. This is what our Bayesian, time-varying beta, stochastic volatility (henceforth BTVBSV) approach accomplishes. Stochastic, time-varying betas have been recently found to be crucial

economic risk variables and past information variables, and ARCH-style conditional betas à la Bollerslev, Engle and Wooldridge (1988). The results in their paper are unaffected by selecting simple, Fama-MacBeth style five-year rolling OLS regression betas and therefore represent our baseline, classical approach to the implementation of MFAPMs.

ingredients of conditional asset pricing, in the sense that there is a growing evidence that careful modelling of the dynamics in factor exposures may provide a decisive contribution to solve the typical anomalies associated with unconditional implementations of multi-factor models. For instance, Jostova and Philipov (2005) find that in the typical Fama and MacBeth's style exercise, the CAPM is rejected with rolling OLS beta estimates while the opposite verdict emerges when they allow for stochastic variation (in the form of a simple AR(1) process) in the conditional CAPM betas.[8] Similarly, Ang and Chen (2007) show that the persistence in the betas help explain the book-to-market effect in the cross section of stock returns. In practice, we specify the relationship between excess returns and factors and the time-varying dynamics in factor loadings and idiosyncratic volatility in the following state-space form

$$\begin{aligned} r_{i,t} &= \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \sigma_{it}\epsilon_{i,t} \\ \beta_{ij,t} &= \beta_{ij,t-1} + \kappa_{1ij,t}\eta_{ij,t} \qquad j = 0, ..., K, \\ \ln(\sigma^2_{i,t}) &= \ln(\sigma^2_{i,t-1}) + \kappa_{2i,t}\upsilon_{i,t} \qquad i = 1, ..., N, \end{aligned} \tag{4}$$

where $\boldsymbol{\epsilon}_t \equiv (\epsilon_{1,t}, \epsilon_{2,t}, ..., \epsilon_{N,t})' \sim N(0, \mathbf{I}_N)$, $\boldsymbol{\eta}_{i,t} \equiv (\eta_{0,t}, \eta_{1,t}, ..., \eta_{K,t}, \upsilon_{i,t})' \sim N(0, \mathbf{Q}_i)$ with $\mathbf{Q}_i$ a diagonal matrix characterized by the parameters q^2_{i0}, q^2_{i1}, ..., q^2_{iK}, q^2_{iv}. Stochastic variations (breaks) in the level of both the beta coefficients and of the idiosyncratic variance σ^2_{it} are introduced and modelled through a mixture innovation approach as in Ravazzolo, Paap, van Dijk and Franses (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 and—for the sake of simplicity—these are assumed to be uncorrelated among one another (i.e., across assets as well as factors) and over time.

This specification is very flexible as it allows for both constant and time-varying parameters. When $\kappa_{1ij,\tau} = \kappa_{2i,\tau} = 0$ for some $t = \tau$, then (4) reduces to (1) when the factor loadings and the quantity of idiosyncratic risk are assumed to be constant, as $\beta_{ij,\tau} = \beta_{ij,\tau-1}$ and $\ln \sigma^2_{i,\tau} = \ln \sigma^2_{i,\tau-1}$. However, when $\kappa_{1ij,\tau} = 1$ and/or $\kappa_{2i,\tau} = 1$, then a break hits either beta or idiosyncratic variance or both, according to the random walk dynamics $\beta_{ij,\tau} = \beta_{ij,\tau-1} + \eta_{ij,\tau}$ and $\ln(\sigma^2_{i,\tau}) = \ln(\sigma^2_{i,\tau-1}) + \upsilon_{i,t}$ (or $\sigma^2_{i,\tau} = \sigma^2_{i,\tau-1}\exp(\upsilon_{i,\tau})$). Note that because when a break affects the betas and/or variances, the random shift is measured by variables collected in $\boldsymbol{\eta}_{i,t}$, we can also interpret $\mathbf{Q}$ not only as a standard, "cold" measure of the covariance matrix of the random breaks in $\boldsymbol{\eta}_{i,t}$, but also of the "size" of such breaks: a large q^2_j means for instance that—whenever $\beta_{ij,t}$ is hit by a break—such a shift is more likely to be large (in absolute value). This process for factor loadings and idiosyncratic residual risk is considerably more plausible than structures typical of the time-varying parameter asset pricing literature in which factor loadings are assumed to vary continuously (i.e., in every period) and usually according to simplistic AR(1) structures with high persistence and small variance for the shocks, such as $\beta_{ij,t} = \phi_{1ij}\beta_{ij,t-1} + \eta_{ij,t}$ and $\ln(\sigma^2_{i,t}) = \phi_{2i}\ln(\sigma^2_{i,t-1})$ $+\upsilon_{i,t}$ with ϕ_{1ij} and ϕ_{2j} close to but less than one.[9]

[8] Jostova and Philipov (2005) point out that OLS betas allow a resarcher to easily estimate the average level of systematic risk but they do not track very well its time patterns, especially when the true but unknown betas are very persistent.

[9] The likelihood tends to be not well-behaved when ϕ_{1ij} and ϕ_{2i} are close to one and their estimation might be difficult, see the discussion and examples in De Pooter, Ravazzolo, Segers and van Dijk (2008).

We estimate (4) using a Bayesian approach, which is in fact the only numerically feasible estimation method for a model with the features of our BTVBSV framework.[10] Realistic values for the different prior distributions obviously depend on the problem at hand. In general, we use weak priors, excluding the size of the breaks $\mathbf{Q}_i$ and the probabilities $\Pr(\kappa_{1ij,\tau} = 1)$ and $\Pr(\kappa_{2i,\tau} = 1)$ for which our priors are quite informative. All other priors imply that the posteriors tend to be centered around their maximum likelihood estimates which eases comparisons with the existing literature.[11] Once estimates *of the posterior density* for unknown coefficients are obtained, we also implement a second-stage estimation pass by estimating, for each month, the following cross-sectional multivariate regression:

$$r_{i,t} = \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1} + e_{i,t} \qquad i = 1, ..., N, \tag{5}$$

where $e_{i,t} \sim N(0, \sigma_t^2)$ and $\beta_{ij,t|t-1}$ measures the *expected* time t sensitivity of asset i to factor j, based on all information available on and upon to time $t-1$. $\beta_{ij,t|t-1}$ is carefully constructed for the purposes of our investigation: it is obtained by taking the lagged value from the updating step of the Kalman filter (see the Appendix for details) and simulating the occurrence of future breaks and the shock magnitude from the appropriate posteriors. This is the exact analog of the logic that has advised Ferson and Harvey (1991) to estimate (3) using one-month lagged value of $\hat{\beta}_{ij,t-1}^{60}$: time t excess return on asset i should be determined by investors with reference only to information available up to time $t-1$ but keeping into account all features of the model (4) known up to time $t-1$.

Even though our Bayesian estimation approach is still articulated on two distinct steps, it is built to avoid all kinds of generated regressor problems that have plagued applications of Fama-MacBeth's methods so far. The second pass estimation is performed similarly to Ouysse and Kohn (2010) to overcome the notorious error-in-variables problem that plagues traditional empirical MFAPMs in small samples. In fact, our Bayesian approach provides an elegant way to take into account estimation uncertainty by averaging out over parameter draws. In particular, we consider the full posterior distribution of the expected factor sensitivities $\beta_{ij,t|t-1}$: for each draw of the betas at a given time t, corresponding values for the risk premia are "drawn" from the relevant posterior distribution; then for each time t we obtain the entire empirical distribution of a given large number of draws of $\lambda_{j,t}$ on which to base our multivariate inferences on. To avoid generated regressor problems in the most resolute form, for each time t we avoid collapsing the posterior density of the factor loadings $\beta_{ij,t|t-1}$ to a single value (e.g., their mean or median) and use instead the entire posterior for the betas. In practice, we draw a large number of times from such a posterior across all N assets and for each draw we estimate a multivariate cross-sectional regression to obtain a corresponding (implicit) draw for the risk premia. As a result, S draws from the joint posterior of the $\beta_{ij,t|t-1}$ will generate S implicit (estimated by OLS regression) draws for the risk premia $\lambda_{j,t}$, which

[10] For instance, in classical MLE framework it would be hard to separately identify the stochastic shifts represented by the variables $\kappa_{1ij,t}$ and $\kappa_{2i,t}$ from the continuous shocks in $\eta_{ij,t}$ and $\upsilon_{i,t}$ without specifying a parametric process for $\kappa_{1ij,t}$ and $\kappa_{2i,t}$, which is however undesirable. In a Bayesian framework, proposing plausible priors informed by economic principles greatly helps to deal with these issues.

[11] These priors are commonly referred to as uninformative or "flat". However, Section 5.3 reports results obtained using tighter, more informative priors and show that these have a negligible impact on our qualitative findings.

represents their posterior joint density.

2.4. *Decomposition Tests*

Independently of the estimation methods employed, we use the estimated time series of factor loadings and risk premia from Sections 2.2-2.3 to perform a number of decomposition exercises. In the case of our BTVBSV approach, of course we shall be obtaining posterior densities for the time series of factor loadings and risk premia, and not simply time series of point estimates. In particular, (3)-(5) decompose excess asset returns in each time period in a component related to risk, represented by the term $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$ plus a residual $\lambda_{0,t}+e_{i,t}$.[12] In principle, a multi-factor model is as good as its implied percentage of total variation in excess returns explained by the first component, $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$. However, here we should recall that even though (3)-(5) refer to excess returns, these are simply statistical implementations of the asset pricing framework in (1). This implies that in practice it may be excessive to expect that $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$ be able to explain most (or even much) of the variability in excess returns. A more sensible goal seems to be that $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$ 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 instrumental variables that proxy for available information at time $t-1$, $\mathbf{Z}_{t-1}$,

$$r_{i,t} = \theta_{i0} + \sum_{m=1}^{M} \theta_{im} Z_{m,t-1} + \xi_{i,t}, \tag{6}$$

to compute the sample variance of the resulting fitted values,

$$Var[P(r_{it}|\mathbf{Z}_{t-1})] \equiv Var\left[\hat{\theta}_{i0} + \sum_{m=1}^{M} \hat{\theta}_{im} Z_{m,t-1}\right], \tag{7}$$

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

$$\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1} = \theta'_{i0} + \sum_{m=1}^{M} \theta'_{im} Z_{m,t-1} + \xi'_{i,t} \tag{8}$$

to compute the sample variance of fitted risk compensations:

$$Var\left[P\left(\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right] \equiv Var\left[\hat{\theta}'_{i0} + \sum_{m=1}^{M} \hat{\theta}'_{im} Z_{m,t-1}\right]. \tag{9}$$

The predictable component of excess returns in (6) 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:

$$Var\left[P\left(r_{i,t} - \hat{\theta}_{i0} - \sum_{m=1}^{M} \hat{\theta}_{im} Z_{m,t-1}|\mathbf{Z}_{t-1}\right)\right] \equiv Var\left[\hat{\xi}_{i,t}\right].$$

[12] Because there are slight differences in the notations used in (3) vs. (5), even though both models have the same economic meaning, in what follows notations from (5) are adopted without any loss of generality. For instance, $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$ is written as $\sum_{j=1}^{K} \hat{\lambda}_{j,t}\hat{\beta}^{60}_{ij,t-1}$ in the standard Fama-MacBeth implementation.

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

$$VR1 \equiv \frac{Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var[P(r_{it}|\mathbf{Z}_{t-1})]} > 0 \tag{10}$$

$$VR2 \equiv \frac{Var\left[P\left(r_{i,t}-\hat{\theta}_{i0}-\sum_{m=1}^{M}\hat{\theta}_{im}Z_{m,t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var[P(r_{it}|\mathbf{Z}_{t-1})]} > 0. \tag{11}$$

VR1 should be equal to 1 if the multi-factor model is correctly specified, which means that all the predictable variation in excess returns ought to be captured by variation in risk compensations; at the same time, VR2 should be equal to zero if the multi-factor model is correctly specified.[13] Notice that $VR1 = 1$ does not imply that $VR2 = 0$ and viceversa, because

$$\begin{aligned} Var[P(r_{it}|\mathbf{Z}_{t-1})] &= Var\left[\hat{\theta}_{i0}+\sum_{m=1}^{M}\hat{\theta}_{im}Z_{m,t-1}\right] \\ &\neq Var\left[P\left(\sum_{j=1}^{K}\hat{\lambda}_{j,t}\hat{\beta}_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]+Var\left[P\left(r_{i,t}\text{-}\hat{\theta}_{i0}\text{-}\sum_{m=1}^{M}\hat{\theta}_{im}Z_{m,t-1}|\mathbf{Z}_{t-1}\right)\right] \end{aligned} \tag{12}$$

Therefore testing whether either $VR1 = 1$ or $VR2 = 0$ or both, remains a sensible strategy in logical terms.

Importantly, when these decomposition tests are implemented using the estimation outputs obtained for our BTVBSV framework, we preserve complete consistency with our Bayesian framework: drawing from the joint posterior densities of the factor loadings $\beta_{ij,t|t-1}$ 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 becomes possible to actually compute VR1 and VR2 in correspondence to each of such draws. This means that any large set S of draws from the (matching) posterior distributions for the $\{\beta_{ij,t|t-1}\}$ and $\{\lambda_{j,t}\}$ generates a posterior distribution for the statistics $VR1$ and $VR2$ as well. This makes it possible to conduct standard Bayesian "inferences" concerning the properties of $VR1$ and $VR2$ in our sample.

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

$$\begin{aligned} Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right] &= \sum_{j=1}^{K}Var\left[P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]+ \\ &+\sum_{j=1}^{K}\sum_{k=1}^{K}Cov[P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right), P\left(\lambda_{k,t}\beta_{ik,t|t-1}|\mathbf{Z}_{t-1}\right)] \end{aligned} \tag{13}$$

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

[13] Ferson and Harvey (1991) claim that this is too strong a condition to be met by any model. Therefore it cannot be interpreted too strictly.

because of the existence of the latter term, the equality

$$\sum_{j=1}^{K} \frac{Var\left[P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]} = 1 \tag{14}$$

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.[14]

3. Data and Summary Statistics

Our paper is based on a large number of monthly time series (31) sampled over the period 1979:12 - 2008:07 for a total of 344 observations per series. The series belong to three main categories. The first group, "Portfolio Returns", includes several asset classes like stocks, bonds and real estate, organized in portfolios, a procedure that is useful to tame the contribution of non-diversifiable risk. The stocks are publicly traded firms listed on the NYSE, AMEX and Nasdaq (from CRSP) and sorted according to two criteria. First, we form 10 industry portfolios by sorting firms according to their four-digit SIC code. Second, we form 10 additional portfolios by sorting (at the end of every year, and recursively updating this sorting in every year in our sample period) NYSE, AMEX and Nasdaq stocks according to their size, as measured by aggregate market value of the company's equity. Using industry and size-sorted criteria to form spread portfolios of stocks to trade-off "spread" and reduction of idiosyncratic risk due to portfolio formation, is typical in the empirical finance 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 different and non-overlapping information on the underlying factors and risk premia.

Data on long-term (10-year) and medium-term (5-year) government bond returns are from Ibbotson and available from the CRSP return tapes. 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 the CRSP return tapes. 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's (1979) approximation formula.[15] Finally, data on REIT total returns come from the North American Real Estate Investment Trust (NAREIT) Association and consists of data on three major categories of tax-qualified REITs, i.e. equity, mortgage, and hybrid equity/mortgage REITs using breakdowns common in the literature.[16] We also use the Wilshire Real

[14] The fact that in (1) the risk factors are assumed to be orthogonal does not imply that their time-varying total risk compensations ($\lambda_{j,t}\beta_{ij,t|t-1}$ for $j = 1, ..., K$) should be orthogonal. For instance, unit risk premia on many factors are well-known to co-move during the business cycle and very precise counter-balancing moves in risk exposures will be required for the covariance term in (13) to equal zero.

[15] Approximated returns from this formula are correlated with actual, Baa rating bracket returns (from Bloomberg) over recent years (2005-2008), with a correlation in excess of 0.8.

[16] REITs (Real Estate Investment Trusts), estabilished in the United States in 1960, constitute a way to invest into large scale, income producing real estate. They have a unique structure: among the requirements to meet, at least 75% of their assets has to be invested in real estate and at least 90% of their taxable earnings have to be paid out as dividends. A popular view, backed by some academic work during the 1990s (e.g., see Karolyi and Sanders, 1998), represents REITs either as mixture of equity- and bond-like features or as equity products similar in many respects to small capitalization stocks. Of course, the two views are hardly consistent with each other.

Estate Securities Index for equity and hybrid REITs and real estate operating companies (office, retail, industrial, apartment, and miscellaneous). All excess return series are computed as the difference between total returns and 1-month T-bill returns, as usual.

Finally, we use a range of macroeconomic factors as standard proxies for the systematic, economy-wide risk factors potentially priced in asset returns. Lagged values of these risk factors (or simple transformation of the factors) are also used as "instruments" when relevant in our methodology, our logic being that all these variables belonged to the information set of the investors when they had made their portfolio decisions. In practice, we employ seven factors (as in Ling and Naranjo, 1997): the excess return on a wide, value-weighted market portfolio (r_t^M) that includes all stocks traded on the NYSE, AMEX, and Nasdaq (from CRSP); the trailing, 12-month dividend yield on all stocks traded on the NYSE, AMEX, and Nasdaq (computed from CRSP data); the default risk premium (def_t) measured as the difference between Baa Moody's *yields* and yields on 10-year government bonds; the change in the term premium ($\Delta term_t$), the difference between 5-year and 1-month Treasury yields; the unexpected inflation rate ($UInfl_t$), computed as the residual of a simple ARIMA(0,1,1) model applied to (seasonally adjusted) CPI inflation; the rate of growth of (seasonally adjusted) industrial production (IP_t); the rate of growth of (seasonally adjusted) real personal consumption growth (PC_t); the 1-month real T-bill rate of return computed as the difference between the 1-month T-bill nominal return and realized CPI inflation rate (not seasonally adjusted).

Table 2 presents summary statistics for the time series under investigation. Note that the table concerns the shorter period 1983:01 - 2008:07 (307 observations per time series) although data for the 1979:12 - 1982:07 data were actually available. There are three reasons for this choice. First, in a portion of our estimation experiments, we use this 3-year period to compute the priors that investors were likely to hold as of the beginning of 1983 based on their recent experience in the difficult financial markets of the early 1980s. Second, the classical Fama-MacBeth methodology ends up requiring the loss of initial observations to initialize the estimation of time-varying betas. Third, 1983:01 is the starting sample data in an important benchmark paper, Karolyi and Sanders (1998), who have also jointly analyzed the properties of U.S. stock, bond, and equity portfolios within a MFAPM. To preserve maximum consistency with their experiments and results, we have chosen to approximately align our initial sample date to theirs. In fact, to favor additional comparability, Table 2 also presents summary statistics for two different sub-samples, 1983:01 - 1992:12 and 1993:01 - 2008:07. In particular, the table reports sample means, sample standard deviations, and the resulting Sharpe ratios (computed with reference to 1-month T-bill returns).

None of the summary statistics in Table 2 is surprising. Most industry portfolios and all cap-sorted portfolios have mean returns between 12 and 14% per year in the overall sample period. Moreover, for all stock portfolios (but one, energy stocks) median returns are substantially higher than mean returns, which is a clear indication of asymmetric return distributions (negative skewness). Volatilities tend to be between 14 and 24 percent in annualized terms, which is also to be expected. Here precise patterns appear, in the sense that small stock portfolios generally tend to be more volatile than large stocks, while the most volatile industries are high tech and durable goods stocks, which may indeed be the most exposed sectors in terms of business cycle dependence. As a result, most Sharpe ratios are in the 0.1-0.17 range (on a monthly basis), with very few outliers such as high tech, durable goods (with ratios below 0.1) and non-durable goods with a Sharpe ratio close to 0.2. There is nothing special to report with reference to returns on 5- and 10-year government bonds, apart from their stunning Sharpe ratios in excess of most stock portfolios, due to the fact that our sample is dominated by the disinflation and declining interest rates of the early 1980s. The Moody's Baa corporate bond portfolio is characterized by mean and median returns that exceed, as they should, those of government bonds free of default risk, although such extra return seems to be compensating their volatility with an exceptional monthly Sharpe ratio of 0.26. The summary performance statistics for real estate portfolios contain instead some unexpected information. While equity REITs are characterized by means (12% per year), volatility (13%), and a Sharpe ratio (0.18) directly comparable to those of stocks (for instance, the value-weighted CRSP portfolio has a mean return of 12%, volatility of 14.8% and a Sharpe ratio of 0.14), mortgage and hybrid REITs have produced much lower mean returns (around 5% per year) but display volatilities in excess of long-term bonds, with resulting Sharpe ratios close to zero. However, because for most of our sample overall REIT portfolios such as NAREIT composite and DJ Wilshire US REIT have come to be dominated by equity REITs, the result is that the corresponding Sharpe ratios are generally close to those of the stock market indices.

The second and third groups of statistics in Table 2 concern the shorter 1983-1992 sample. As for 1983-1992, the summary statistics can be compared to Karolyi and Sanders' (1998) and in general we fail to notice any significant differences. However, the differences between sub-periods are considerable. For instance, and possibly contrary to common perception, it is evident that the post-1992 age has been a rather disappointing period for stocks, and this emerges independently of the portfolio sorting criterion employed, with the only exception of small and energy stocks. Even though a few volatilities are lower in the post-1992 periods than in the earlier sub-sample, the generalized decline in mean stock returns implies lower Sharpe ratios in the 1993-2008 period. This is also reflected in the statistics concerning the market portfolio, which has recently yielded lower mean returns (9% vs. 15%), imposed higher volatility on investors (15% vs. 14%), and has yielded as a result a Sharpe ratio of 0.10 vs. 0.18 in the 1983-1992 period. These conclusions are reversed when it comes to real estate returns: for four of the portfolios analyzed, the post-1992 Sharpe ratio exceeds the 1983-1992 ratios; for instance the ratio for NAREIT composite jumps to 0.16 from 0.09 in the earlier period, and a solid contribution is given by mortgage REITs which had negative Sharpe ratios in the 1980s and switch to positive ratios in the past 10 years. Finally, some instability characterizes our return data for long-term bonds returns: these display significantly less volatility in the post-1992 period but also lower mean returns. As a result, their Sharpe ratios decline from an exceptionally high range of 0.23-0.36 over the 1983-1992 period to 0.13-0.17 in the post-1992 sample.

4. **Empirical Results**

4.1. *Factor Loadings*

Figures 1-2 plot a selection (see below) of the time series of 5-year rolling window estimates of the betas in (1), obtained from a classical two-pass Fama-MacBeth approach. To save space, Figures 1-2 report plots of time series of rolling window betas for all the 28 portfolios used in our estimation, but for two specific (potential) risk factors only: the U.S. market portfolio and the term premium (the slope of the riskless yield curve).[17] Other, similar Figures concerning the remaining risk factors—the default spread, the real short term rate, the IP growth rate, the unemployment rate, and real consumption growth—are available upon request even though we summarize their contents and implications below. Clearly—unless one is ready to make very restrictive assumptions—the two-pass Fama-MacBeth does not allow us to compute and report standard error bands around the estimated coefficients.[18]

Figure 1 shows that all the 20 equity-like portfolios yield a positive and high beta on the value-weighted market portfolio, with rather limited dynamics over the sample (the only exception seems to be utility stocks, whose market beta declines towards zero during the 1980s and 1990s and then jumps up after 2000). As one would expect, high-tech and, to some extent, stocks in the two lowest size deciles are aggressive portfolios, with their market betas often exceeding 1 and appear to have somewhat increased over time. Non durable, energy, utility, and to some extent decile 10 stocks are instead defensive and yield market betas at or below 1. The three bond portfolios are characterized by negligible or even slightly negative market betas. REITs appear instead to be a mixture between equities and bonds, as their market betas are generally positive but also centered around a value of 0.5. However, most real estate indices also display a visible upward trend that becomes strong in the late 1990s. Figure 2, with reference to term premium exposures gives a general impression of instability in betas. For most equities, the mid-1990s and recently the post-2005 period are all characterized by large, positive, and increasing exposures to term structure risk. Although most equity portfolios start with zero or even slightly negative term structure betas, by the end of our sample durables, high tech, retail, and the first two size deciles (positive and large), and utilities (negative) display large betas in absolute value. If we take the term premium to be a business cycle indicator—in the sense that a higher (lower) term premium signals an improvement (deterioration) of business cycle conditions, see e.g. Estrella and Hardouvelis (1991)—these results for industrial and size-sorted stock portfolios are sensible. Both medium- and long-term Treasury bond excess returns have a negative exposure to the term premium, while corporate bonds generally do not. Real estate index returns display strong time-variation in term spread betas, with a massive downward trend (starting from 0–0.5 to -4 or -5, in the case of equity REITs) between 2001 and 2006. However, the yield curve betas recover ground towards zero or even the positive range between 2006 and 2008.

Results (unreported) for the credit risk premium show that instability also dominates the exposures of U.S. traded assets to default risk. The betas for all 20 equity portfolios change sign at least three times over our sample period. In many cases (e.g., non durables, high-tech, utilities, size deciles 1-4 and 10) the range of variation of the default premium beta equals or exceeds 8, with betas that typically start out positive and large (between 3 and 6) decline and turn negative by the late 1990s, and then increase

[17] These two risk factors are selected because we perceive a widespread belief in the asset pricing literature that variables such as market portfolio returns and the term structure of interest rates may heavily affect excess returns on all assets and portfolios. As the paper will show, this may actually not be the case in the light of the empirical evidence.

[18] Reporting standard (asymptotic) normal-based standard error bands would be valid only assuming that the beta coefficients are stable over time, which contradicts the nature of the entire experiment.

again after the turn of the millennium. Since a variety of papers (see e.g., Fama and French, 1989, as well as Stock and Watson, 2003) have argued that a surging default risk premium is a predictor of economic downturns, it is sensible to find that durables, manufacturing, retail, and the first three size decile stocks have the highest exposures to default premium-type business cycle risk, while non-durables, utility, health and decile 10 stocks have essentially zero average exposures to this risk. Government bonds have a positive an relatively stable exposure, while Baa corporate bonds tend to show more variation, with betas turning negative in the late 1990s. A positive exposure of Treasury returns is quite interesting, because it means that unexpected increases in the credit risk premia ought to cause—net of the contemporaneous effects of the business cycle on other variables—an increase in government security prices. This is a "flight to quality" effect in which the credit spread may increase not only (or not even mainly) because Baa yields increase, but also because Aaa yields decline. Real estate indices that include any equity components yield wildly gyrating exposures, oscillating between -2 and +7 and with peaks in correspondence to the early 1980s and mid-1990s. As one would expect given the nature of the underlying business, the NAREIT mortgage index generates the maximal exposures in absolute value but oscillates over time.

First-stage rolling estimates of the loadings on the level of the (real, ex-post) Treasury bill rate imply that, as stressed by a number of recent papers (e.g., see Ang, Piazzesi and Wei, 2006), it may be often advisable to use not only the slope of the yield curve but also some measure of level—for instance as measured by the T-bill rate—to capture the dynamics of the business cycle. In this case, we observe tremendous variability over time as well as heterogeneity across industries. In the presence of such wild swings in betas any analyst would be really best advised to refrain from expressing any economic interpretation. It is difficult to make sense of exposures that are at the same time close to +50 in some industries while in others they are approaching -50, not to mention the enormous loading provided by a beta of 50 (in absolute value), even after discounting the modest variability of changes in real T-bill rates. In all equity portfolios there are dramatic reversals with betas changing sign between the late 1990s and 2002. In some industries (e.g., non durables, retail) the exposure to real rate shocks plummets and then bounces back to almost zero after 2005. For other portfolios (e.g., high-tech, utilities, the intermediate size deciles) the opposite occurs. We are at a loss at trying to tell a story for what type of economic phenomena may cause such a pervasive and puzzling gyrations. If we take the real short-term rate as a business cycle indicator—for instance, because the Fed tries to push the real rate down (up) reacting to weakening (over-heating) economic conditions, according to a Taylor rule—it is hard to imagine why high-tech and utility stocks may present similarly signed exposures and trends over time. It is even more shocking to find positive exposures of government bond excess returns to real short term rates during the period 1998-2002, as increasing rates should lead to losses and negative excess returns in bond portfolios. Finally, REIT portfolios (with the only exception of hybrids) all have negative and relatively stable exposures to this factor.

Finally, the betas on economic factors that directly link to business cycle conditions—IP growth, the change in the unemployment rate, and real consumption growth—all reveal exposures that generally wildly oscillate around zero although their averages tend to be slightly positive, as if pure measures of the level of real economic activity may have played at best a minor role over our sample. This of course related to the typical and puzzling finding of a disconnect between asset returns and real business cycle conditions, observed by a vast literature. However, the IP growth betas of mortgage and hybrid excess REIT returns are volatile but generally negative, which is sensible if interest rate effects—an economic upturn generally causes an increase in long-term mortgage rates—are more important determinants of property values than the state of the business cycle.

Figures 3-9 report instead medians and 90% Bayesian credibility intervals computed from the posterior densities of the loadings $\beta_{ij,t}$ over time from the BTVBSV model. The plots are structured in the same way as Figures 1-2, to favor comparisons. However, Figure 5-9 report results for a subset of 16 portfolios only (8 industry stock portfolios, and all bond and real estate portfolios) for reasons of space. An overview of the plots reveals immediately that the Bayesian estimates of the loadings are considerably smoother than the rolling window ones. This is a first interesting result: even though (4) formally allows the $\beta_{ij,t}$ to be subject to jumps over time, as a result of the realization of a latent binary random variable, the resulting posterior densities are actually smoother than what one could retrieve using a naïve rolling window scheme. Interestingly, this smoothness mimics exactly what many earlier papers have imposed by assuming near unit root processes ($\beta_{ij,t} = \phi_{ij}\beta_{ij,t-1} + u_{ij,t}$) with small variance of the shocks, but is derived endogenously, as required by the data, which means that occasional large jumps in exposure and/or high volatility of the corresponding process may be accommodated.[19] Second, with a limited number of exceptions that will noted below (mostly concerning market betas), the 90% confidence bands are generally wide and often end up including zero, which means that the betas are estimated with very little reliability.

In particular, Figure 3 collects most of the loadings for which we have evidence they are not zero. All equity portfolios have positive betas, although none of them display reliably estimated trends or variability in exposures.[20] However, four portfolios (non durables, energy, health, utilities) yields betas with an upper 90% band that is always below 1, an indication that these are equity portfolios that are reliably estimated to be defensive; two portfolios (high-tech and size decile 2) yields betas with an upper 90% band that is always above 1, an indication that these are equity portfolios that are reliably estimated to be aggressive. These findings confirm similar results reported with reference to first-stage Fama-MacBeth betas. None of the bond portfolios gives evidence of non-zero exposure to the market portfolio, even when such exposures are precisely estimated. Similarly to what reported in Figure 1, real estate portfolios confirm their defensive nature—the corresponding upper 90% bound rarely includes 1, although their lower 90% bound is also never touching zero—which may be also interpreted as resulting from some hybrid nature, between stocks and bonds. In Figure 4, concerning the term premium betas, we start getting the first strong differences between the implications of BTVBSV and the classical Fama-MacBeth approach: most equity portfolios fail to be significantly exposed to term premium risk, in the sense that their 90% bands systematically include zero.[21] Here the naive two-stage Fama-MacBeth approach may have proven misleading, inducing

[19] For instance in Figure 3 the high-tech beta on the market portfolio jumps by non-negligible amounts a few times.

[20] This means that for most of the 20 equity portfolios it is always possible to draw a straight horizontal line that is everywhere contained in the 90% credibility bands. The only exceptions are size deciles 9 and 10 but this occurs not because of the presence of important trends, but because the 90% bands are narrow and closely track the trendless variation in the betas.

[21] The solitary exception is given by high-tech stocks, which yield a positive exposure with a posterior median centered

a researcher into assessing large yield curve slope exposures for durables, high tech, retail, and small caps (positive) and utilities (negative) that do not appear to be there when any instability in betas and residual variances are formally incorporated in the analysis. It is even more striking to note that in Figure 4 none of the government bond portfolios yields non-zero or accurately estimated term premium exposures, once the effects of the market betas are netted out. Finally, there is some residual evidence of real estate exposure to the yield curve slope factor, and as in Figure 2 the corresponding betas are positive and small for mortgage REITs, and negative and small for equity REITs. All in all, it seems that explicitly accommodating the instability in exposures through (4) may substantially weaken the popular belief that the term structure of interest rates is an important business cycle factor driving realized excess asset returns.

Figure 5 concerns instead credit risk premium betas and provides further confirmation of the differences induced by the adoption of (4). With minor exceptions, none of the equity portfolios carries a reliably estimated exposure to the default factor. In the case of the 10 industry portfolios, the betas are mostly flat with posterior medians centered around zero. In this case, also bond and real estate portfolios do not have any important exposures: only mortgage REITs and (over the period 1983-1997) equity REITs show positive betas with a lower 90% bound that fails to include zero. Therefore our earlier finding that a number of equity portfolios (e.g., durables, manufacturing, retail, and size decile 1-3) and especially Treasury bonds had significant exposure to a credit risk factor seem to vanish when instability in betas is formally modelled and not only naively captured by rolling window estimates. However, the finding of some positive exposure of real estate to this factor seems robust. Figure 6 confirms these findings with reference to changes in the level of the real short rate, the only marginally exception being that the smallest, first-decile stocks are characterized by a small negative beta whose posterior does not include zero.

Figures 7-9 give a homogeneous message as far the exposures of equity portfolios are concerned: for each of them, either the real activity factors—IP growth, changes in the unemployment rate, and personal consumption growth—never mattered, in the sense that their 90% posterior bands always included a zero beta exposure, or they have stopped mattering over time, in general after 1989. Consider for instance Figure 7: for all the stock portfolios presented (but the same holds for the remaining 12 equity portfolios not in the picture), either the 90% bands always include a zero beta, or (this is the case for non durables and retail stocks, that start with negative exposures) the bands widen to include zero after the late 1980s. The picture is more complex when it comes to bond and real estate portfolios. However the implication remains simple: IP growth matters quite a bit for long-term bonds (both Treasuries and corporate junk) and real estate; changes in the unemployment rate and the rate of personal consumption growth do not.

4.2. *Where There Any Persistent Misspricings?*

Figures 10 and 11 report estimates of $\beta_{i0,t}$. In an ICAPM interpretation of (1) and under the null of correct specification, the time series $\{\beta_{i0,t}\}$ gives indications on time-varying misspricing. If $r_{i,t}$ represents returns in excess of the riskless rate and the K factors have been correctly specified, $\beta_{i0,t} \neq 0$ represents evidence of non-zero excess returns for a portfolio i with zero exposures to the K risk factors, which implies the

between 0.5 and 1.

existence of an arbitrage opportunity and it is inconsistent with first principles (e.g., non-satiation). In the finance literature, estimates of quantities like $\beta_{i0,t}$ are often named Jensen's alphas (see Jensen, 1968) and interpreted as measures of abnormal (excess) returns, where their abnormality refers to the principle that only systematic risk ought to be priced.

Figure 10 starts by presenting 5-year rolling window estimates obtained from Fama-MacBeth first stage regressions. In each of the 28 panels we also report the mean of the $\hat{\beta}_{i0,t}$s over time and the p-value of a test of the null $E[\beta_{i0,t}] = 0$ using the sample mean as an estimator and HAC-corrected standard errors to remove the effects of the MA(59) structure induced by the 5-year rolling window estimation scheme. Figure 10, for the two-step Fama-MacBeth case, reports a rather incredible finding: all 28 portfolios display large and negative estimates of $\beta_{i0,t}$ for most of the sample. The means over the entire 1983-2008 sample range from -2.6% per month in the case of energy stocks and -8.5% in the case of high tech stocks. Even taking into account the fact that the estimates are largely overlapping, all the means appear to have been precisely estimated and negative. The $\hat{\beta}_{i0,t}$s start out at very low values (sometimes close to a stunning -20% per month) during the early 1980s, increase to zero between the mid-1990s and the 2000s (depending on the portfolio considered), and in some cases (for most equity portfolios and all the real estate ones) become positive between the late 1990s and 2005. Between 2006 and 2008, as it is reasonable to expect in the light of the recent financial crisis, all $\hat{\beta}_{i0,t}$s decline again and all become negative. This set of implications for misspricings from the standard Fama-MacBeth procedure is clearly implausible: on average, over a rather long 25-year period, all the 28 portfolios would have been significantly over-priced, and as a result their abnormal returns would have been negative.[22] The under-pricing *of the risk exposures* derives from the fact that—when their time-varying, rolling window risk exposures are taken into account—the 28 portfolios appear to have yielded excess returns that are meager with respect to what the model would have required. The over-pricing appears to have been substantial, although a considerable portions of the $\bar{\beta}_{i0}$ reported in Figure 10 derive from the large, negative abnormal returns from 1983-1986. From a statistical perspective, it is also difficult to forget that many time series of rolling window betas turned out to be extremely volatile, almost erratic, and that in many situations it has been hard to provide any intuitions to the patterns revealed by Figures 1-2. Finally, note that while in the last three decades the literature has often debated the existence of bubbles in portions of the U.S. capital markets (e.g., in the technology industry in the late 1990s, in the long-term government bond market in 2005-2006, and in the real estate market in 2004-2007), the two-pass estimation procedure points to a situation of generalized, persistent and significant over-pricing of all publicly traded assets in the U.S., which seems rather implausible in the light of a widespread faith in the efficiency of U.S. capital markets.

Although a reaction to the poor performance of the model would advise to revise the structure of the MFAPM—for instance, by increasing the number of factors, or changing their identity—the goal of

[22] Note that when looking at excess realized returns, a negative sequence of $\hat{\beta}_{i0,t}$ over time implies under-pricing of risk because this means that historical excess returns have been lower than what their risk exposures imply. However, a negative $\hat{\beta}_{i0,t}$ at time t implies a low expected excess returns and hence (for instance, in a simple risk-adjusted present value model) over-pricing of the asset at time t. The two perspectives are logically consistent because under-pricing the risk exposures of an asset implies that the asset will be over-priced and will fail to yield adequate rates of capital gain and pay-outs over time.

this paper is to show that by revising the method of implementation of the MFAPM, and by explicitly modelling the existence of instability in risk exposures and in the level of idiosyncratic risk, a much more satisfactory outcome may be obtained. Figure 11 reports posterior medians and 90% credibility intervals for the $\beta_{i0,t}$s estimated from (4). For comparability, we report in each panel the average of the median posteriors over time, although strictly speaking such a statistic has no precise inferential meanings. First, out of 28 portfolios, in 5 cases we find $\beta_{i0,t}$s with a posterior median that is uniformly positive over our entire sample period (this occurs for non durables, energy, telecommunication, health, and utility stocks), although there are other 14 portfolios (durables, high tech, retail, size deciles 1-6 stocks, as well as long-term government bonds and all the NAREIT portfolios) for which the posterior means of the abnormal returns are negative. In the remaining 9 cases, the mean $\beta_{i0,t}$ changes sign over time. In essence, the values of the posterior means of the $\beta_{i0,t}$s as well as their sign are economically plausible. Even more interestingly, once we take into account of the 90% confidence bands, it is clear that in 28 out of 28 portfolios, the reported posterior means fail to be precisely estimated, in the sense that the 90% band always includes a zero abnormal return throughout the entire period 1983-2008. Although this might be premature, one might say that the posterior means of the $\beta_{i0,t}$s fail to give a reason to think that the model is inadequate.[23]

4.3. *Idiosyncratic Risk*

A growing literature (see e.g., Campbell, Lettau, Malkiel and Xu, 2001, Brandt, Brav, Graham and Kumar, 2008, and Zhang, 2010) has stressed that the idiosyncratic variance of the excess returns of most sets of test portfolios, σ^2_{it}, has undergone interesting shifts and/or dynamics over the last two decades. Figure 12 displays plots from the classical, two-pass Fama-MacBeth method. The presence of rich dynamics is obvious for all the portfolios. Two patterns emerge. First, idiosyncratic variance tends to be large for most portfolios investigated, a sign that the two-pass method provides a very rough explanation to the 28 time series of excess returns data. To make this impression compelling, in each of the panels in Figure 12 we have reported the ratio between the average of the two-pass estimate $\overline{\hat{\sigma}^2}_{it}$ and the sample variance of portfolio returns in Table 1, over the period 1983-2008. Note that because of the rolling window nature of the $\{\hat{\sigma}_{it}\}$, this ratio may actually exceed one. Anything close to 1 or in excess of 1 of course denotes a very poor average fit of the MFAPM. We note that a large fraction (14) of the 28 portfolios investigated yields a ratio between $\overline{\hat{\sigma}^2}_{it}$ and the sample variance that is close to or in excess of 0.5, with all the bond portfolios with ratios in excess of one, and all the real estate portfolios with ratios between 0.6 and 0.8. In this cases, the non-systematic component of excess returns still explains at least 50% (and often exceeds 100%, meaning that the factor model in some sense clouds our understanding) of the total variance of excess returns. For instance, an analysis of the fitted excess returns from the MFAPM shows that the variations in expected excess returns hardly correspond to any of the realized variations in excess returns. Second, most equity portfolios (in practice, all the industry portfolios and size deciles 1-8) record a peak

[23] Figure 11 shows that durable stocks (especially after 2000), high-tech stocks (throughout the entire sample), and size deciles 1-6, and mortgage REITs (especially after 2000) have yielded excess returns substantially below the levels justified by their risk exposures and therefore can be considered to have been over-priced.

in idiosyncratic variance between 2000 and 2003. In some cases, the rolling window idiosyncratic variance practically doubles between 1999 and 2004, meaning that the MFAPM loses most of its ability to fit excess returns using risk exposures.[24] This is partially visible also in the case of the real estate portfolios, although here the trend for idiosyncratic risk to increase seem to have continued after 2004. This is consistent with the evidence in Campbell, Lettau, Malkiel and Xu (2001) and Zhang (2010) who report a positive trend in idiosyncratic volatility during the 1962-2001 period (while aggregate market volatility remained constant) that has strongly reverted from 2001 onwards to start climbing up again since 2007.

Figure 13 reports instead the posterior medians for σ_{it}^2 estimated from (4), along with 50% credibility intervals.[25] For all equity portfolios it is always possible to draw a horizontal line that fits within the 50% confidence bands (being wider, this clearly applies to 90% bands as well). This implies that when all the uncertainty is taken into account, there is no strong evidence of trends in idiosyncratic variance, even though the plots for the individual portfolios show some evidence of a peak in the early 2000s and some sign of a growing trend towards the end of our sample. The ratios between posterior medians for the σ_{it}^2 and sample variances for excess returns, are generally similar to those found in the case of Figure 12, although the specific values differ. The BTVBSV model performs poorly for energy and health stocks, while it explains away almost all the variability in excess returns in the case of medium and large cap stocks. The most important differences between classical two-step Fama-MacBeth results and the Bayesian results appear for the bond and REIT portfolios. While long-term government bonds are now characterized by a idiosyncratic variance comparable to equity portfolios (as a fraction of sample variance), medium-term Treasuries practically do not have any idiosyncratic variance. This means that the Bayesian estimates of (4) are severely tilted to fit 5-year Treasury excess returns.[26] However, this is not the case for Baa corporate bonds, for which the ratio in Figure 13 is 1.54, only modestly below the 1.99 in Figure 12.

4.4. *Risk Premia*

Figure 14 reports second-pass Fama-MacBeth time series of estimates of the risk premia $\{\hat{\lambda}_{j,t}\}$ $(j = 1, ..., K)$ from (3). In the figure, 8 panels appear, seven for the macro risk factors assumed in the paper, and one for $\hat{\lambda}_{0,t}$ which can be interpreted as a measure of overall, cross-sectional abnormal performance. Also in this case, as in Figures 1-2, the instability in the estimates is overwhelming, with most $\hat{\lambda}_{j,t}$ series either frequently switching sign or being even centered around a general, negative sample mean, which means negative premia for positive risk exposures. In Figure 14, we also report in the boxes within each plot, the time series sample mean of each of the risk premia as well as the p-value of a test of the null of $E[\lambda_{j,t}] = 0$ using the sample mean as an estimator and HAC-corrected standard errors. None of the risk premium series turns out to be estimated with any statistical precision, apart from the risk premium on the (change in the) yield curve slope factor, for which the average $\hat{\lambda}_{j,t}$ is -0.19 with a p-value between 1 and 5%. However,

[24]Some equity portfolios (e.g., manufacturing, shops, middle cap stocks) also display a first peak in correspondence to the mid-1980s.

[25]In the case of idiosyncratic variance the 90% bands would be so large to completely obscure any time variation in the posterior median for the purposes of visualization.

[26]The same applies to size deciles 7-10 stock portfolios.

such a premium is then significantly negative, which—assuming that a steeper yield curve forecasts an improvement in business cycle conditions—is puzzling. Moreover, the average risk premium on market risk turns out to be imprecisely estimated, and to be small (0.28). The only quantity for which there is compelling evidence is in fact $\hat{\lambda}_{0,t}$ which is estimated to be negative and massive (-3.92% per month), with an essentially zero p-value. In a sense, all that a standard estimation Fama-MacBeth approach to (1) reveals is that U.S. asset return data contain strong evidence of structural misspricings in the form of negative abnormal returns, which is consistent with our remarks on Figure 10.

The Bayesian (posterior median) estimates in Figure 15 are more stable and—more importantly—a few of them are precisely estimated. This is only superficially paradoxical: when a model that incorporates instability is adopted, stable estimates are obtained. However, it is clear that a better (correctly) specified model offers efficiency gains and impresses a stability in the resulting estimates that may be responsible for this finding. In particular, the market risk premium has an average (over time) posterior median of 0.232 that is often significant (i.e., the corresponding 90% confidence band fails to contain zero) in 56% of our sample. A careful inspection of Figure 15 reveals that there is some persistence in market risk premium medians, with long periods over which this premium is either significantly negative (1984-85, 1987-88 after the October '87 crash, late 1992-1993, and 2001-2002) or positive (e.g., 1996-1997 and again 2004-2006). Figure 15 also shows that two other risk premia—on default and IP growth risk—can be estimated with some precision, at least over sub-samples (e.g., 1993-1995 and 2005-2006 for the default risk premium and 2003-2005 for IP growth). However, on average both risk premia have posterior medians that are very small, 0.018 in the case of default risk and 0.056 in the case of IP growth.

Table 2 reports summary statistics for median posteriors of the $\lambda_{j,t}$s, compared with those for the second-pass point estimates, the $\hat{\lambda}_{j,t}$s, from Fama-MacBeth. In this case, we also report the empirical (over the entire 1983-2008 period) 5th and 95th percentiles of each series, besides the sample standard error for the average of the $\lambda_{j,t}$s (note that using them for inference implicitly embraces a frequentist perspective). To gain additional insights, we split our 25-year long sample in two, 1983-1992 and 1993-2008. Once more, the classical estimation procedure that non-parametrically tracks time-variation in the parameters using 5-year rolling window estimates delivers nonsensical insights. Over the full sample, only one factor seems to be precisely priced in U.S. asset returns—this is default/credit risk—but this seems to happen with the wrong sign, -0.19. If we realistically assume that an increase in default risk spreads forecasts business cycle downturns, the result is that publicly traded assets would imply a negative price for business cycle risk. During the first 1983-1992 sub-sample, the results make even less sense because not only default risk would command a reliably estimated and negative price (-0.35), but the same now occurs also for IP growth risk (-0.30). However, during the longer and final sample 1993-2008, volatility takes over and none of the risk premia is reliably estimated. On the top of these issues, the fact that the $\hat{\lambda}_{0,t}$s are significant, large, and negative re-emerges throughout the three samples: U.S. financial markets would have been characterized by massive under-pricing of risk and hence over-pricing of all asset classes.

Fortunately, a much more comforting picture emerges from the Bayesian estimation exercise, when a parametric model for unstable exposures and idiosyncratic risk are imposed. Here one result is striking:

consistently with earlier evidence centered on real estate data (see e.g., Ling and Naranjo, 1997) in the overall sample, real consumption growth is the only precisely estimated (p-value is 0.030) risk premium, it displays the expected positive sign, even though the size of the premium tends to be small (0.06). Also the Bayesian design gives evidence of moderate but precisely estimated misspricings, with an average posterior median of $\lambda_{0,t}$ equal to 0.37% per month. However, these misspricings now correspond to over-pricing of the risk factors captured by the model and to a structural but mild under-pricing of the assets over time. Also in this case, there is evidence of a better performance of the MFAPM in the first sub-sample than in the second: during the 1983-1992 period, the average posterior median of $\lambda_{0,t}$ converges to zero (0.26%) and fails to be significant. Three factors are significantly priced, two—the term spread and changes in real short term rates (i.e., in some ways both level and slope of the yield curve)—with the expected signs (0.51 and 0.09, respectively), while unemployment risk carries a negative sign (-0.07, with p-value between 5 and 10%) that however is not completely surprising in the light of some asset pricing literature (see e.g., Boyd, Hu and Jagannathan, 2005). During the second sub-sample, the finding of an accurately estimated over-pricing of the risk factors (0.45%) and of real consumption growth being the only accurately estimated risk premium (now 0.11 with a p-value below 1%) re-appears. This evidence is surprisingly tilted towards concluding that U.S. assets are priced by business cycle-related factors—real consumption growth as in the C-CAPM paradigm, but with periods in which also riskless yield curve factor seems to play an important role—than towards the typical finding that the cross-section of excess returns would depend on other financial factors, in this case the value-weighted market portfolio factor. Moreover, the fact that at least one sub-sample could be found for which $\lambda_{0,t}$ does not seem "significant" in a traditional sense—along with the evidence in Figure 11 that very few of the $\beta_{i0,t}$ posterior series would yield sample medians estimated with any precision—may be taken as an encouraging first step towards failing to reject the BTVBSV model.

5. Economic Tests

So far our discussion has focussed on the statistical performance of the models with emphasis on whether there was evidence of either the $\lambda_{0,t}$s or the $\beta_{i0,t}$s coefficients being different from zero. We have concluded that (1)-(2) is rejected in its standard, two-pass Fama-MacBeth implementation based on 5-year rolling window estimates. However, there was encouraging evidence that the BTVBSV model may be less at odds with the data. Therefore, with reference to (4), we compute the VR1 and VR2 ratios in Section 5.1. In Section 5.2 we factor $Var[P(\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1})]$ as the sum of the contributions given by each of the factors, leaving the covariance terms as a residual interaction effect. In both Sections, the information at time $t-1$ ($\mathbf{Z}_{t-1}$) is proxied by the instrumental variables listed in Section 3, plus a dummy variable to account for the "January effect".

5.1. *Variance Ratios*

The first two columns of Table 3 present posterior medians of $VR1$ and $VR2$ obtained from (4) for each of the 28 portfolios under examination. Variance ratio results are encouraging, although there is some difference between the $VR1$ vs. the $VR2$ perspectives. Under a VR1 perspective, we can claim that approximately half of the predictable variation in excess returns is indeed captured by our MFAPM. Such percentages are in fact very high and satisfactory for the size-sorted stock portfolios, more heterogeneous (but again, generally encouraging) for the industry equity portfolios, mediocre for REIT portfolios, and rather low when it comes to bond portfolios. In the case of size-sorted portfolios, $VR1$ always exceeds 50% and it averages close to 70%, which is quite impressive. In the case of industry portfolios, one goes from peaks in excess of 70% for high tech and retail stocks, to a minimum of 22% in the case of utility stocks, even though most portfolios comfortably exceed 50%. Also the REIT portfolios yield VR1 statistics around 50%. Finally, it is well know from work by Cochrane and Piazzesi (2005 and following papers) that—albeit certainly present in the data—the predictability in excess bond returns (at least as far as riskless bonds are concerned) remains hard to pin down using macroeconomic factors and appears instead to be driven by the implicit dynamics of the riskless yield curve, as captured by special (tent-like) patterns implicit in forward rates.

However, because $VR1 + VR2 = 1$ does not hold, the finding of good VR1 ratios fails to imply that the VR2 ratios are as close to zero as much as we would want, which is consistent with the finding in Figure 13 of large and (to some extent) time-varying idiosyncratic variances. In 19 portfolios out of 28 we at least find that VR2 is below 50%. VR2 is indeed uniformly moderate (below 0.4, with the only exception of very small caps) for the size-sorted equity portfolios. However, it is much less satisfactory for the REIT and bond portfolios, for which there is evidence that large chunks of predictable variation is actually picked up by the dynamics in idiosyncratic risk.

5.2. *Decomposing Predictable Variation*

Table 3 shows that the predictable variation in excess stock returns is mostly explained by the market risk factor: with one exception (the smallest capitalization portfolio), all the ratios $Var[P(\lambda_{MKT,t}\beta_{iMKT,t|t-1}| \mathbf{Z}_{t-1})]\ /Var[P(\sum_{j=1}^{7}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1})]$ exceed 0.6 with peaks in excess of 0.8 for a number of industries as well as the medium- and large-cap portfolios.[27] This result on the importance of the market portfolio to explain predictable variation in stock excess returns stands in contrast to our statistical findings in Section 4. The origin of such findings is soon obvious: the market factor explains little or nothing of the predictable variation in excess bond returns. REIT portfolios stand in between, with a contribution of the market risk factor between 0.19 and 0.47. As far as stocks are concerned, the next most important factor contributions come from IP growth and changes in the real short term rate, although the heterogeneity across portfolios is strong and mostly concentrated among industry portfolios. In the case of bond portfolios, many other factors—e.g., default and yield curve slope risks, but to some extent also business cycle factors such as unemployment and the real short term rate—different from the market give contributions to explain the predictable variation in excess returns. For instance, the (large) predictable portion of long-term bond

[27] As explained in Section 2.4, these ratios may exceed 100% because $Var[P(\sum_{j=1}^{7}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1})]$ will also reflect the contribution of covariance terms between factor terms. In fact, in Table 3 the only two contributions exceeding 100% are obtained in the presence of sizably negative covariance contributions.

excess returns seems to be mostly driven by default risk (representing a likely flight-to-quality effect during economic downturns), the yield curve spread factor, and the real short term rate. Once more, real estate assets fall in between with important contributions from the credit risk factor as well as the market. Finally, in spite its statistical importance, the contribution of the real consumption growth tends to be modest. This is not inconsistent with our earlier findings because even though the posterior median series for the consumption risk premium reveals its entity with great precision, this has no implication for the behavior of the exposures to such a risk, which we have seen to be very imprecisely estimated in Figure 9.

6. **Robustness Checks**

Because our research design, especially in its Bayesian implementation, has implied a number of choices at several steps, in this Section we conclude with a few remarks on the main qualitative differences when estimation has been repeated after changing any of the choices made. In general we have found modest differences and this provides further support to the encouraging results revealed by Section 4 and 5. To save space, we have not plotted or tabulated complete set of results, that remain available from the Authors.

6.1. *Bayesian Two-Stage Fama-MacBeth Strategies*

As explained in Section 2.3, our Bayesian implementation follows Ouysse and Kohn (2010) as a way to side-step all kinds of generated regressor problems that have plagued applications of Fama-MacBeth's methods. In practice, this means that we have considered the full posterior distribution of the time $t-1$ (expected) factor sensitivities $\beta_{ij,t|t-1}$: for each draw of the betas at time t, corresponding values for the risk premia are "drawn" from the relevant posterior distribution; then, again at time t, we obtain the entire empirical distribution of a given large number of draws of $\lambda_{j,t}$ on which to base our multivariate inferences. In spite of its internal consistency and the fact this appears to be the only way in which all relevant (estimation) uncertainty may be taken into account, to see whether our economic insights may be mostly driven by this innovative choice, we have repeated the entire analysis by using the simpler but traditional approach of using plain (median, in our case) betas from the first pass rather than their full posterior distribution at time t. Moreover, for each draw s of the risk premia ($\lambda_{j,t}$), we compute their average over time ($\widehat{\lambda}_j^i$) in order to obtain their empirical distribution. Equivalently, this is the same as stating that we have constructed a time series of estimated risk premia for each iteration s of the Bayesian algorithm and consequently computed the average risk premia, for which a posterior distribution can then be easily obtained.[28]

Because this robustness exercise is a compromise between the advantages of the Bayesian approach that allows the estimation of complex models such as (4) and the classic two-step approach, in this case we have reported this mixed "Bayesian/Fama-MacBeth" results in Table 2. Even though the posterior

[28] We have also tried to simply use median betas from the first pass rather than their full posterior distribution at time t leaving the portion of the algorithm concernint the $\lambda_{j,t}$ intact. While there are no sharp differences in terms of uncertainty surrounding the $\lambda_{j,t}$ for each t as revealed by their posterior distributions, the results are quite different as far as the dynamics of such posterior distributions over time is concerned. The increased dispersion in the estimates is striking and would prevent us from drawing any of the implications currently in Section 4.4.

median estimates are never dramatically different from those obtained in the "correctly estimated" Bayesian implementation (e.g., the posterior medians switch sign in only 5 cases out of 24), there are a few important changes in the inferential implications vs. the previous two methods in the other panels of Table 2. First, during the second, more recent sub-sample, the MFAPM is a complete "wash-out" in the sense that none of the risk premia is precisely estimated and presumably different from zero, including $\lambda_{0,t}$. On the one hand, a λ_0 centered at zero is good news for a MFAPM; on the other hand, the fact that none of the risk premia is precisely estimated means that the model has little to say. During the two remaining periods, it is interesting to note that default, term structure, real short rate, and real consumption growth risks all stop being significantly priced, while the market risk factor is, with p-values always below 1%. However, during the full sample it remains the case that even this mixed implementation reveals substantial over-pricing of risks, in the sense that the average $\lambda_{0,t}$ is 0.38% and precisely estimated, so that the chances for this model to hold throughout our sample are identical to those discussed in Section 4.4.[29]

6.2. *Informative Priors*

Finally, we have experimented with an informative prior in the second pass in order to put some structure (constraints) on the distribution and moments of the risk premia. These are now postulated to be normally distributed with zero mean and variance such that there is 95% probability that annualized premia are smaller in absolute value than the maximum return observed in the sample for all the assets.[30] We record a striking reduction in the variability of the estimated posterior distributions (as well as their medians) for the risk premia relative to the baseline case. Although the qualitative results and insights from Figure 15 and Table 2 apply intact, there is one important change: the market risk premium becomes now "significant" in both the full sample and in the first sub-sample, 1983-1992. In essence, using informative priors on the premia which constraints their variability, we find both less variable premia (so far the result has been built in the type of prior used) and economic implications that encompass the second and third panels of Table 2: both real consumption growth and market risks are important drivers of the cross section of U.S. returns, even though during the initial samples other factors (term spread, unemployment, and the real short term rate) appear to have been priced. However, we have decided to present the results of this further exercise as a mere robustness check because it should be clear that the selection of the priors exerts a first-order effect on our conclusions on whether the classical CAPM-style market risk factors are priced.

7. Conclusions

We have analyzed and compared the empirical performance of two alternative ways in which a standard MFAPM with time-varying risk exposures and premia may be estimated. The first method echoes the two-pass approach advocated by Fama and MacBeth (1973) used in a substantive body of applied work in empirical finance. However, as it is well known, such an two-stage approach is plagued by difficult problems

[29]In fact, the 90% confidence bands for $\lambda_{0,t}$ stop including zero also for the 1993-2008 sample, although a classical-style t-test yields a p-value of 0.157.

[30]A complete description of prior distributions and hyperparameters used can be found in the Appendix.

with errors-in-variables and arbitrariness of the choice of the rolling windows. The second approach extends previous work by Ouysse and Kohn (2010) and is based on a formal modelling of the latent process followed by risk exposures and idiosyncratic volatility capable to capture structural shifts in parameters.

Our application to monthly, 1979-2008 U.S. data for stock, bond, and publicly traded real estate returns shows that the classical, two-stage approach that relies on a nonparametric, rolling window modelling of time-varying betas yields results that are unreasonable. For instance, there is evidence that all the portfolios of stocks, bonds, and REITs examined in this paper would have been grossly over-priced during our sample period, which is a rather bizarre result inconsistent with any faith in the efficiency of U.S. capital markets. Moreover, very few risk factors appear to be priced and, when they are, they carry the wrong sign (in the sense that more business cycle exposure ought to lead to negative risk premia). On the contrary, the empirical implications of our Bayesian estimation of (4) are plausible and there are indications that the model may be consistent with the data. For instance, most portfolios do not appear to have been grossly misspriced and a few risk premia are precisely estimated with a plausible sign. Real consumption growth risk turns out to be the only factor that is persistently priced throughout the sample. However, we cannot claim to have achieved complete success: the BTVBSV ends up giving an acceptable empirical performance only "on the shoulders" of what is a dwarf, in the sense that the Fama-MacBeth methodology leads to a disappointing fit. It would be interesting both to further fine-tune the standard, more traditional part of the model—such as the number of factors to be specified as well as their nature and definition—and at the same time to work on the specific structure and assumptions appearing in (4) to test whether its empirical performance may be improved and/or any different insights may be derived.

References

[1] Ang A. and Chen J., "CAPM Over the Long Run: 1926-2001", *Journal of Empirical Finance*, 2007, 1-40.

[2] Ang A., Piazzesi M. and Wei M., "What Does the Yield Curve Tell us About GDP Growth?", *Journal of Econometrics*, 2006, 131, 359-403.

[3] Boyd J. H., Hu J. and Jagannathan R., "The Stock Market's Reaction to Unemployment News: Why Bad News Is Usually Good for Stocks", *Journal of Finance*, 2005, 60, 649-672.

[4] Bollerslev T., Engle R.F. and Wooldridge J.M., "A Capital Asset Pricing Model with Time-Varying Covariances", *Journal of Political Economy*, 1988, 96, 116-131.

[5] Brandt M., Brav A., Graham J. and Kumar A., "The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes?", 2008, *Review of Financial Studies*, forthcoming.

[6] Campbell J.Y., Lettau M. , Malkiel B.G. and Xu Y., "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk", *Journal of Finance*, 2001, 56, 1-43.

[7] Campbell J.Y. and Cochrane J.H, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior", *Journal of Political Economy*, 1999, 107, 205-251.

[8] Chan K.C., Henderhott P. and Sanders A. B., "Risk and Return on Real Estate: Evidence from Equity REITS", *Real Estate Economics*, 1990, 18, 4, 431–452.

[9] Carter C. and Kohn R., "On the Gibbs sampling for state-space models", *Biometrika*, 1994, 81, 541-553.

[10] Chen NF., Roll R. and Ross S., "Economic Forces and the Stock Market", *Journal of Business*, 1986, 59, 383–403.

[11] Cochrane J.H., *Asset Pricing*, Princeton University Press, Princenton: NJ, 2005.

[12] Cochrane J.H. and Piazzesi M., "Bond Risk Premia", *American Economic Review*, 2005, 95, 138-160.

[13] De Pooter M., Ravazzolo F., Segers R. and van Dijk H.K., "Bayesian Near-Boundary Analysis in Basic Macroeconomic Time Series Models", *Advances in Econometrics*, 2008, 23, 331-402.

[14] Dittmar R.F., "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns", *Journal of Finance*, 57, 369-403.

[15] Estrella, A., Hardouvelis, G.A., "The Term Structure as Predictor of Real Economic Activity", *Journal of Finance*, 1991, 46, 555–576.

[16] Fama E. and French K.R., "Business Conditions and Expected Returns on Stocks and Bonds", *Journal of Financial Economics*, 1993, 81, 607-636.

[17] Fama E. and MacBeth J., "Risk, Return and Equilibrium: Empirical Tests", *Journal of Political Economy*, 1973, 81, 607-636.

[18] Ferson W. and Harvey C. R., "The Variation of Economic Risk Premiums", *Journal of Political Economy*, 1991, 99, 385-415.

[19] Ferson W. and Schadt R., "Measuring Fund Strategy and Performance in Changing Economic Conditions", *Journal of Finance*, 1996, 51, 425-461.

[20] Gerlach R., Carter C. and Kohn R., "Efficient Bayesian Inference for Dynamic Mixture Models", *Journal of the American Statistical Association*, 2000, 95, 819-828.

[21] Giordani P., Kohn R. and van Dijk D., "A Unified Approach to Nonlinearity, Outliers and Structural Breaks", *Journal of Econometrics*, 2007, 137, 112-133.

[22] Giordani P. and Kohn R., "Efficient Bayesian Inference for Multiple Change-Point and Mixture Innovation Models", *Journal of Business and Economic Statistics*, 2008, 26, 1, 66-77.

[23] Groen J.J.J., Paap R. and Ravazzolo F., "Real-time Inflation Forecasting in a Changing World", *Federal Reserve of New York Staff Reports*, 2009, 388.

[24] Hofmann, M., Gatu, C., Kontoghiorghes, E.J., "Efficient Algorithms for Computing the Best Subset Regression Models for Large-Scale Problems", *Computational Statistics and Data Analysis*, 2007, 52, 16-29.

[25] Ilmanen, A., "Time-Varying Expected Returns in International Bond Markets", *Journal of Finance*, 1995, 50, 481-506.

[26] Jensen, M.C., "The Performance of Mutual Funds in the Period 1945-1964", *Journal of Finance*, 1968, 23, 389-416.

[27] Jostova G. and Philipov A., "Bayesian Analysis of Stochastic Betas", *Journal of Financial and Quantitative Analysis*, 2005, 40, 4, 747-778.

[28] Karolyi G.A. and Sanders A., "The Variation of Economic Risk Premiums in Real Estate Returns", *Journal of Real Estate Finance and Economics*, 1998, 17, 3, 245-262.

[29] Kapetanios, G., "Variable Selection in Regression Models Using Nonstandard Optimisation of Information Criteria. *Computational Statistics and Data Analysis*, 2007, 52, 4-15.

[30] Kim S., Shepard N. and Chib S., "Stochastic volatility: likelihood inference and comparison with ARCH models", *Review of Economic Studies*, 1998, 65, 361-93.

[31] Koop G., *Bayesian Econometrics*, Wiley, 2003.

[32] Ling, D.C. and Naranjo A., "Economic Risk Factors and Commercial Real Estate Returns", *Journal of Real Estate Finance and Economics*, 1997, 15, 3, 283–307.

[33] Liu C. H. and Mei J., "The Predictability of Returns on Equity REITs and their Co-Movement with Other Assets", *Journal of Real Estate Finance and Economics*, 1992, 5, 401-418.

[34] Merton R., "An Intertemporal Capital Asset Pricing Model", *Econometrica*, 1973, 41, 867-887.

[35] Ouysse R. and Kohn R., "Bayesian Variable Selection and Model Averaging in the Arbitrage Pricing Theory Model ", *Computational Statistics and Data Analysis*, 2010, 54, 12, 3249-3268.

[36] Primiceri G., "Time Varying Structural Vector Autoregressions and Monetary Policy", *Review of Economic Studies*, 2005, 72, 821-852.

[37] Raftery A.E. and Lewis S.M., "How Many Iterations in the Gibbs Sampler?", Bernardo J.M (eds), *Bayesian Statistics 4*, Oxford University Press, 763-773.

[38] Ravazzolo F., Paap R., van Dijk D. and Franses P.H., "Bayesian Model Averaging in the Presence of Sructural Breaks", Rapach D. and Wohar M. (eds.), "Forecasting in the Presence of Structural Breaks and Uncertainty", *Frontiers of Economics and Globalization*, Elsevier, 2007.

[39] Ross S., "The Arbitrage Theory of Capital Asset Pricing" *Journal of Economic Theory*, 1976, 13, 341–360

[40] Shiller R.J., "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure", *Journal of Political Economy*, 1979, 87, 1190-1219.

[41] Stock J.H. and Watson M.W., "Forecasting Output and Inflation: The Role of Asset Prices", *Journal of Economic Literature*, 2003, 41, 788-829.

[42] Tanner M.A. and Wong W.H., "The Calculation of Posterior Distributions by Data Augmentation", *Journal of the American Statistical Association*, 1987, 82, 528-550.

[43] Zhang C., "A Re-examination of the Causes of Time-varying Stock Return Volatilities", *Journal of Financial and Quantitative Analysis*, 2010, 45, 663-684.

Appendix

We separately present each of the steps of our Bayesian implementation of the two-step Fama-MacBeth (1973) approach.

First pass:

For each asset $i = 1, ..., N$, the model in (4) is

$$\begin{aligned} r_{i,t} &= \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \sigma_{i,t}\epsilon_{i,t} \\ \beta_{ij,t} &= \beta_{ij,t-1} + k_{ij,t}\eta_{ij,t} \qquad j = 0, ..., K \\ \ln(\sigma^2_{i,t}) &= \ln(\sigma^2_{i,t-1}) + k_{2i,t}\upsilon_{i,t} \qquad i = 1, ..., N, \end{aligned}$$

where $\boldsymbol{\epsilon}_t \equiv (\epsilon_{1,t}, \epsilon_{2,t}, ..., \epsilon_{N,t})' \sim N(0, \mathbf{I}_N)$, $\boldsymbol{\eta}_{i,t} \equiv (\eta_{0,t}, \eta_{1,t}, ..., \eta_{K,t}, \upsilon_{i,t})' \sim N(0, \mathbf{Q})$ with $\mathbf{Q}$ a diagonal matrix characterized by the parameters $q_0^2, q_1^2, ..., q_K^2, q_\upsilon^2$, and $\boldsymbol{\kappa}_t \equiv (\kappa_{0,t}, ..., \kappa_{K,t}, k_{2,t})'$ is a $((K+2)\times 1)$ vector of unobserved uncorrelated 0/1 processes with $\Pr[\kappa_{j,t} = 1] = \pi_j$ for $j = 0, ..., K+1$ and $\Pr[k_{2,t} = 1] = \pi_{2k}$. The model parameters are the structural break probabilities $\boldsymbol{\pi} \equiv (\pi_0, ..., \pi_K, \pi_2)'$ and the vector of variances of the break magnitude $\mathbf{q}^2 \equiv (q_0^2, q_1^2, ..., q_K^2, q_\upsilon^2)$. They are collected in a $(2(K+1)\times 1)$ vector $\boldsymbol{\theta} \equiv (\boldsymbol{\pi}', (\mathbf{q}^2)')'$.

Independent conjugate priors are used to ease posterior simulation. For the break probability we assume simple Beta distributions,

$$\pi_j \sim Beta(a_j, b_j) \qquad \pi_j^{k_2} \sim Beta(a_j^{k_2}, b_j^{k_2}), \tag{15}$$

where the hyperparameters a_j and b_j $(j = 0, ...K+1)$ reflect prior beliefs about the occurrence of breaks. For the variance parameters the inverted Gamma-2 prior is chosen,

$$q_j^2 \sim IG(\nu_j, \delta_j) \qquad q_\upsilon^2 \sim IG(\nu_\upsilon, \delta_\upsilon), \tag{16}$$

where ν_j expresses the strength of the prior mean.

For posterior simulation we run the Gibbs sampler in combination with the data augmentation technique by Tanner and Wong (1987). The latent variables $B = \{\beta_t\}_{t=1}^T$, $R = \{\sigma_t^2\}_{t=1}^T$, and $\mathcal{K} = \{\kappa_t\}_{t=1}^T$ are simulated alongside the model parameters, $\boldsymbol{\theta}$. The complete data likelihood function is given by

$$\begin{aligned} p(r, B, \mathcal{K}, R|\boldsymbol{\theta}, F) \quad &= \quad \prod_{t=1}^{T} p(r_t|F_t, \beta_t, \sigma_t^2) \prod_{j=0}^{m} p(\beta_{jt}|\beta_{jt-1}, \kappa_{jt}, q_j^2) \times \\ & \quad\quad p(\sigma_t^2|\sigma_{t-1}^2, k_{2,t}, q_v^2) \prod_{j=0}^{k} \pi_j^{k_{jt}}(1-\pi_j)^{1-k_{jt}} \pi_{2k}^{k_{2t}}(1-\pi_{2k})^{1-k_{2t}}. \end{aligned}$$

Combining the prior and the data likelihood, we obtain the posterior density

$$p(\theta, B, \mathcal{K}, R|r, F) \propto p(\theta)p(r, B, \mathcal{K}, R|\theta, F). \tag{17}$$

Defining $\mathcal{K}_\beta = \{\kappa_{0,t}, ..., \kappa_{K,t}\}_{t=1}^T$ and $\mathcal{K}_\sigma = \{k_{2t}\}_{t=1}^T$, the sampling scheme consists of the following iterative steps:

1. Draw $\mathcal{K}_\beta$ conditional on $R, \mathcal{K}_\sigma, \boldsymbol{\theta}$, and r.
2. Draw B conditional on $R, \mathcal{K}, \boldsymbol{\theta}$ and r.
3. Draw $\mathcal{K}_\sigma$ conditional on $B, \mathcal{K}_\beta, \boldsymbol{\theta}$, and r.
4. Draw R conditional on $B, \mathcal{K}, \boldsymbol{\theta}$ and r.
5. Draw $\boldsymbol{\theta}$ conditional on $B, \mathcal{K}$ and r.

The first step applies the efficient sampling algorithm of Gerlach, Carter and Kohn (2000), the main advantage being drawing $\kappa_{j,t}$ without conditioning on the states $\beta_{j,t}$, as Carter and Kohn (1994) instead do. The conditional posterior density for $\kappa_{\beta,t}$, $t = 1, .., T$ unconditional on B is:

$$\begin{aligned} p(\kappa_{\beta,t}|\mathcal{K}_{\beta,-t}, \mathcal{K}_\sigma, R, \boldsymbol{\theta}, r) \quad &\propto \quad p(r|\mathcal{K}_\beta, \mathcal{K}_\sigma, R, \boldsymbol{\theta})p(\kappa_t|\mathcal{K}_{\beta,-t}, \boldsymbol{\theta}) \\ &\propto \quad p(r^{t+1,T}|r^{1,t}, \mathcal{K}, R, \boldsymbol{\theta})p(r_t|r_{1,t-1}, \kappa_{\beta,1,t-1}, R, \boldsymbol{\theta}, x)p(\kappa_{\beta,t}|\mathcal{K}_{\beta,-t}, \boldsymbol{\theta}). \end{aligned} \tag{18}$$

Gerlach, Carter and Kohn (2000) show how to evaluate the first two terms while the last one is obtained from the prior. When $\mathcal{K}_{\beta,t}$ and $\beta_{j,t}$ are highly dependent, the sampler of Carter and Kohn (1994) breaks down completely: the higher the correlation (dependence), the bigger the efficiency gain. The latent process for the betas is estimated by means of the forward-backward algorithm of Carter and Kohn (1994).

$\mathcal{K}_\sigma$ and R are drawn in the same way as $\mathcal{K}_\beta$ and B. To do so we follow Kim, Shepard and Chib (1998) and approximate the log of a $\chi^2(1)$ distribution by means of a mixture of seven normals. In each iteration of the Gibbs sampler we simulate a component of the mixture distribution in order to get a conditional linear state space model for $\ln(\sigma_t^2)$. Finally, the vector of parameters $\boldsymbol{\theta}$ is easily sampled as we use conjugate priors.

We use a burn-in period of 1,000 and draw 5,000 observations storing every other of them to simulate the posterior distributions of parameters and latent variables. The resulting autocorrelations of the draws are very low.[31]

Second pass:

To estimate the cross section in (5) at each time t and for each draw of $B_{t|t-1} = (\beta_{1,t|t-1}, ..., \beta_{N,t|t-1})$ where each $\beta_{j,t|t-1}$ is a $((K+1) \times 1)$ vector and N is the total number of assets, we use natural conjugate priors. In particular,

$$p(\lambda, \sigma^2) = p(\lambda|\sigma^2) \times p(\sigma^2) \tag{19}$$

where

$$(\lambda|\sigma^2) \sim N(\underline{\lambda}, \sigma^2 \underline{V}) \text{ and } (\sigma^2) \sim IG(\frac{\underline{\nu}}{2}\underline{s}^2, \frac{1}{2}\underline{\nu}) \tag{20}$$

Combining them with the data likelihood we obtain a joint posterior density with convenient analytical form. The resulting marginal posterior distributions are

$$\begin{aligned} (\lambda|r) &\sim t(\overline{\lambda}, \overline{s}^2\overline{V}, \overline{\nu}) && (21) \\ (\sigma^2|r) &\sim IG(\frac{\overline{\nu}}{2}\overline{s}^2, \frac{1}{2}\overline{\nu}) && (22) \end{aligned}$$

with

$$\begin{aligned} E(\lambda|r) &= \overline{\lambda} && (23) \\ var(\lambda|r) &= \frac{\overline{\nu s}^2}{\overline{\nu}-2}\overline{V} && (24) \\ E(\sigma^2|r) &= \frac{\overline{\nu s}^2}{\overline{\nu}-2} && (25) \\ var(\sigma^2|r) &= \frac{(\overline{\nu s}^2)^2}{(\overline{\nu}-2)^2(\frac{\overline{\nu}}{2}-2)} && (26) \end{aligned}$$

where

$$\begin{aligned} \overline{V} &= (\underline{V}^{-1} + (X'^{-1})^{-1} && (27) \\ \overline{\lambda} &= (\underline{V}^{-1} + (X'^{-1})^{-1}(\underline{V}^{-1}\underline{\lambda} + (X'^{-1}\hat{\lambda}) && (28) \\ \overline{\nu} &= \underline{\nu} + N && (29) \end{aligned}$$

and $\hat{\lambda}$ is the OLS estimate.[32]

Results are presented with two different sets of priors. In the former case we are noninformative ($\underline{\nu} = 0$ and $\underline{V}^{-1} = 0$) and use the well known Jeffreys' prior while in the latter case we impose some prior information. In more detail, we opted for a small amount of strength ($\underline{\nu} = 5$) supporting a prior view for premiums with zero mean and standard deviation equal to a twelfth of the maximum absolute return observed in the sample. Finally, the prior residual variance is centered at about 10, a value that appeared in the higher range of the maximum likelihood estimates.

[31] In order to gain a rough idea of how well the chain mixes in our algorithm we follow Primiceri (2005) in looking at the autocorrelation function of the draws. Primiceri (2005) plots the 20-h-order sample autocorrelation for some of the parameters. He bases his assessment on two further and more elaborate indicators, such as the inefficiency factor (IF) and the Raftery and Lewis (1992) diagnostic. As for the stochastic variances, we report satisfactorily low levels of autocorrelation that in most cases essentially vanishes after a handful of lags. However, for the long term government bonds we observe a persistent autocorrelation which reaches values close to 0.4 after 20 lags, especially at the end of the sample. It would be interesting to investigate to what extent this lower inefficiency is harmful for inference.

[32] Intermediate steps to derive the marginal posterior distributions have been skept. Interested readers can refer to Koop (2003) for details.

Table 1

Summary Statistics for Financial and Macroeconomic Time Series Used in the Paper

	1983 - 2008				1983 - 1992				1993-2008			
	Mean	Median	Std. Dev.	Sharpe Ratio	Mean	Median	Std. Dev.	Sharpe Ratio	Mean	Median	Std. Dev.	Sharpe Ratio
	10 Industry Portfolios, Value-Weighted (Source: CRSP, NYSE/AMEX/NASDAQ)											
Non Durable Goods	1.240	1.260	4.260	0.195	1.656	1.390	4.686	0.244	0.810	1.100	3.736	0.135
Durable Goods	0.825	1.010	5.956	0.070	1.144	1.045	5.753	0.110	0.496	0.980	6.161	0.031
Manufacturing	1.170	1.410	4.610	0.165	1.325	1.610	4.789	0.170	1.011	1.010	4.426	0.159
Energy	1.314	1.040	5.102	0.177	1.241	1.285	4.567	0.160	1.389	0.920	5.615	0.193
High Tech	0.991	1.150	7.273	0.080	1.034	1.155	5.708	0.092	0.946	1.150	8.616	0.074
Telecommunications	0.938	1.410	5.181	0.102	1.520	1.835	4.393	0.230	0.338	0.570	5.840	0.005
Shops and Retail	1.074	1.270	5.109	0.130	1.329	1.315	5.502	0.149	0.811	1.160	4.672	0.108
Health	1.141	1.200	4.772	0.153	1.485	1.420	5.202	0.187	0.786	0.690	4.271	0.112
Utilities	1.053	1.340	3.926	0.164	1.209	1.145	3.427	0.204	0.892	1.340	4.388	0.133
Other	1.018	1.440	4.776	0.127	1.311	1.830	4.731	0.169	0.715	1.130	4.818	0.085
	10 Size-Sorted Portfolios, Value-Weighted (Source: CRSP, NYSE/AMEX/NASDAQ)											
Decile 1	1.029	1.180	5.762	0.107	0.897	1.170	5.042	0.077	1.165	1.180	6.435	0.133
Decile 2	1.062	1.440	6.018	0.108	1.020	1.650	5.157	0.099	1.105	0.940	6.812	0.117
Decile 3	1.121	1.770	5.579	0.127	1.239	2.350	5.081	0.143	0.999	1.400	6.065	0.114
Decile 4	1.046	1.580	5.439	0.117	1.199	1.920	5.113	0.135	0.887	1.420	5.770	0.101
Decile 5	1.137	1.710	5.296	0.137	1.317	1.995	4.937	0.164	0.951	1.510	5.653	0.114
Decile 6	1.087	1.410	4.876	0.139	1.335	1.615	4.680	0.176	0.831	1.170	5.074	0.103
Decile 7	1.168	1.410	4.739	0.160	1.314	1.240	4.613	0.174	1.018	1.440	4.876	0.146
Decile 8	1.111	1.480	4.803	0.146	1.280	1.730	4.559	0.169	0.936	1.060	5.052	0.125
Decile 9	1.146	1.660	4.343	0.169	1.336	1.740	4.387	0.188	0.949	1.620	4.303	0.149
Decile 10	1.005	1.180	4.223	0.141	1.328	1.485	4.033	0.203	0.670	1.040	4.400	0.083
	Bond Returns (Source: FREDII, Ibbotson via CRSP)											
10-Year Treasury Notes	0.629	0.616	1.180	0.185	0.806	0.826	1.267	0.234	0.446	0.423	1.057	0.132
5-Year Treasury Notes	0.580	0.552	0.771	0.220	0.735	0.738	0.823	0.274	0.419	0.435	0.681	0.165
Baa Corporate Bonds (10-20)	0.824	0.934	1.587	0.261	1.078	1.148	1.582	0.359	0.561	0.582	1.554	0.164
1-month Treasury Bills	0.410	0.410	0.187	0.000	0.510	0.490	0.171	0.000	0.306	0.360	0.141	0.000
	Real Estate Returns (Source: NAREIT and Dow Jones)											
NAREIT - Composite	0.891	1.047	3.687	0.131	0.788	0.655	3.114	0.089	0.998	1.448	4.206	0.164
NAREIT - Equity TR	1.084	1.200	3.790	0.178	1.092	1.033	3.333	0.175	1.075	1.354	4.221	0.182
NAREIT - Mortgage TR	0.448	0.885	5.599	0.007	0.401	0.514	3.744	-0.029	0.497	1.977	7.032	0.027
NAREIT - Hybrid TR	0.437	0.718	5.239	0.005	0.690	0.563	3.679	0.049	0.177	1.031	6.471	-0.020
DJ Wilshire US REIT TR Index	0.976	0.970	3.929	0.144	0.838	0.545	3.430	0.095	1.120	1.478	4.392	0.185
	Economic Risk Variables (Source: FREDII and CRSP, Ibbotson)											
Excess Value-Weigted Market	0.598	1.030	4.253	0.141	0.740	1.030	4.127	0.179	0.451	0.940	4.388	0.103
Default Premium (annualized)	2.093	1.960	0.520	__	1.984	1.920	0.407	__	2.207	2.110	0.596	__
Change in Term Spread	0.000	0.000	0.624	__	-0.010	-0.025	0.728	__	0.010	0.010	0.496	__
Unexpected Inflation	0.009	0.004	0.288	__	0.008	0.006	0.224	__	0.011	0.003	0.342	__
Industrial Production Growth	0.246	0.282	0.532	__	0.285	0.294	0.536	__	0.205	0.199	0.528	__
Real Pers. Consumption Growth	0.228	0.226	0.174	__	0.265	0.243	0.150	__	0.191	0.177	0.190	__
Real 1-month T-Bill Returns	0.145	0.176	0.310	__	0.219	0.217	0.239	__	0.068	0.120	0.354	__
	Instrumental Variables (Source: FREDII, CRSP, Ibbotson via CRSP)											
5Y Govt. Yield - 1m T-Bill (annual)	1.534	1.540	1.092	__	2.028	2.155	1.028	__	1.023	1.000	0.908	__
Yield Spread Baa - Aaa (annualized)	0.553	0.540	0.496	__	0.785	0.790	0.456	__	0.314	0.390	0.417	__
Dividend Yield (annualized)	2.548	2.370	1.017	__	3.417	3.315	0.640	__	1.650	1.690	0.297	__

Table 2

Summary Statistics for Second-Pass Fama-MacBeth and Bayesian Posterior Median Estimates of Risk Premia

	Full sample (1983-2008)						1983-1992						1993-2008					
	Average	**Std. Error**	**t-stat**	**p-value**	**5% CI**	**95% CI**	**Average**	**Std. Error**	**t-stat**	**p-value**	**5% CI**	**95% CI**	**Average**	**Std. Error**	**t-stat**	**p-value**	**5% CI**	**95% CI**
	Classical Two-Step Fama-MacBeth																	
Intercept (Avg. cross-sectional abnormal returns)	-3.922	0.233	-16.826	**0.000**	-10.128	2.857	-4.963	0.385	-12.894	**0.000**	-12.282	0.914	-3.254	0.282	-11.520	**0.000**	-9.187	3.183
Market	0.281	0.265	1.060	0.289	-7.351	7.307	0.369	0.460	0.804	0.421	-6.468	7.406	0.224	0.321	0.699	0.485	-7.548	7.032
Default (credit) spread	-0.190	0.075	-2.518	**0.012**	-2.204	2.008	-0.348	0.116	-3.000	**0.003**	-2.217	1.793	-0.089	0.099	-0.900	0.368	-2.198	2.175
Term spread riskless yields	0.123	0.108	1.145	0.252	-2.797	2.853	0.105	0.203	0.517	0.605	-4.167	3.600	0.135	0.120	1.125	0.261	-2.542	2.646
Unemployment	0.009	0.036	0.243	0.808	-0.845	0.914	0.007	0.051	0.139	0.890	-0.975	0.828	0.010	0.050	0.199	0.843	-0.795	0.988
IP growth	-0.096	0.096	-1.007	0.314	-2.717	2.777	-0.300	0.182	-1.644	**0.100**	-3.433	2.973	0.034	0.104	0.329	0.742	-2.066	2.412
Real consumption growth	0.008	0.029	0.286	0.775	-0.862	0.791	0.015	0.055	0.276	0.782	-0.968	1.115	0.004	0.032	0.119	0.905	-0.616	0.611
Real Treasury Bill	-0.035	0.049	-0.705	0.481	-1.451	1.114	-0.022	0.083	-0.265	0.791	-1.647	1.308	-0.043	0.061	-0.702	0.483	-1.168	0.951
	Bayesian Model with Instability in Loadings and Idiosynchratic Variances																	
Intercept (Avg. cross-sectional abnormal returns)	0.374	0.119	3.145	**0.002**	-2.876	4.017	0.257	0.181	1.416	0.157	-2.733	3.603	0.449	0.157	2.862	**0.004**	-2.916	4.140
Market	0.232	0.264	0.877	0.381	-7.113	7.115	0.395	0.464	0.852	0.394	-6.283	7.841	0.127	0.317	0.401	0.689	-7.113	6.783
Default (credit) spread	0.018	0.107	0.168	0.866	-2.653	2.876	-0.159	0.103	-1.539	0.124	-1.724	1.590	0.131	0.163	0.808	0.419	-3.216	3.747
Term spread riskless yields	0.150	0.137	-1.096	0.273	-3.707	3.787	0.507	0.199	-2.546	**0.011**	-4.000	2.931	0.080	0.183	0.436	0.663	-3.616	4.192
Unemployment	-0.006	0.044	-0.127	0.899	-1.259	1.167	-0.072	0.042	-1.716	**0.086**	-0.768	0.641	0.037	0.068	0.546	0.585	-1.373	1.383
IP growth	0.056	0.109	0.515	0.607	-2.770	3.531	-0.166	0.126	-1.320	0.187	-2.254	2.236	0.199	0.159	1.251	0.211	-2.956	3.855
Real consumption growth	0.059	0.027	2.166	**0.030**	-0.734	0.842	-0.014	0.030	-0.452	0.652	-0.612	0.561	0.106	0.040	2.638	**0.008**	-0.801	1.003
Real Treasury Bill	0.010	0.049	-0.199	0.842	-1.223	1.105	0.090	0.042	2.149	**0.032**	-0.693	0.868	0.074	0.076	-0.974	0.330	-1.447	1.473
	Mixed Bayesian/Fama-MacBeth Approach																	
Intercept (Avg. cross-sectional abnormal returns)	0.376	0.178	2.118	**0.034**	**0.170**	**0.579**	0.258	0.314	0.822	0.411	-0.090	0.590	0.420	0.297	1.416	0.157	**0.082**	**0.786**
Market	0.232	0.087	2.657	**0.008**	**0.020**	**0.436**	0.398	0.123	3.237	**0.001**	**0.090**	**0.700**	-0.075	0.145	-0.516	0.606	-0.470	0.285
Default (credit) spread	0.013	0.219	0.059	0.953	-0.183	0.224	-0.162	0.186	-0.869	0.385	-0.048	0.160	0.145	0.280	0.519	0.604	-0.228	0.540
Term spread riskless yields	0.155	0.243	0.638	0.523	-0.452	0.162	0.514	0.432	1.191	0.234	-1.003	0.007	-0.002	0.407	-0.005	0.996	-0.470	0.539
Unemployment	-0.007	0.305	-0.023	0.982	-0.133	0.116	-0.070	0.886	-0.079	0.937	-0.290	0.134	0.043	0.381	0.113	0.910	-0.170	0.238
IP growth	0.061	0.205	0.297	0.766	-0.172	0.262	-0.160	0.451	-0.355	0.723	-0.530	0.160	0.336	0.320	1.051	0.293	-0.037	0.710
Real consumption growth	0.058	0.467	0.124	0.901	-0.067	0.182	-0.017	0.924	-0.018	0.985	-0.200	0.175	0.127	0.666	0.191	0.849	-0.086	0.354
Real Treasury Bill	-0.005	0.357	-0.014	0.989	-0.154	0.126	0.091	0.935	0.097	0.922	-0.160	0.330	-0.083	0.372	-0.223	0.823	-0.302	0.120

Table 3

Variance Ratio Coefficients and Predictable Variation Decompositions from Bayesian Factor Model with Instability in Risk Exposures and Idiosyncratic Variance

	VR1	VR2	VW	PREM	DSLOPE	UI	IPGRW	CONSUM	REALTB	Int Eff
NoDur	0.48	0.52	0.79	0.02	0.01	0.03	0.12	0.02	0.07	-0.21
Durbl	0.58	0.42	0.74	0.06	0.01	0.02	0.02	0.01	0.03	0.02
Manuf	0.67	0.33	1.35	0.02	0.03	0.07	0.03	0.02	0.03	-0.63
Enrgy	0.54	0.46	0.77	0.03	0.05	0.07	0.15	0.02	0.06	-0.32
HiTec	0.77	0.23	0.85	0.03	0.07	0.01	0.01	0.01	0.02	-0.05
Telcm	0.43	0.57	0.64	0.07	0.04	0.02	0.02	0.02	0.05	0.00
Shops	0.73	0.27	0.67	0.10	0.02	0.01	0.03	0.01	0.02	0.06
Hlth	0.60	0.40	0.80	0.03	0.05	0.02	0.17	0.02	0.24	-0.51
Utils	0.22	0.78	0.70	0.07	0.10	0.06	0.03	0.06	0.14	-0.53
Other	0.52	0.48	0.92	0.01	0.00	0.02	0.02	0.03	0.02	-0.08
CAP1RET	0.53	0.47	0.45	0.05	0.02	0.01	0.01	0.01	0.04	0.32
CAP2RET	0.66	0.34	0.74	0.04	0.00	0.01	0.01	0.01	0.02	0.11
CAP3RET	0.73	0.27	0.83	0.03	0.01	0.01	0.01	0.01	0.02	0.02
CAP4RET	0.71	0.29	0.93	0.06	0.00	0.02	0.01	0.01	0.01	-0.10
CAP5RET	0.73	0.27	0.93	0.06	0.00	0.02	0.01	0.02	0.02	-0.12
CAP6RET	0.71	0.29	0.84	0.07	0.00	0.02	0.01	0.02	0.02	-0.03
CAP7RET	0.71	0.29	0.88	0.02	0.01	0.04	0.02	0.02	0.04	-0.08
CAP8RET	0.67	0.33	0.93	0.02	0.01	0.01	0.02	0.03	0.03	-0.10
CAP9RET	0.60	0.40	1.04	0.01	0.01	0.03	0.04	0.01	0.06	-0.25
CAP10RET	0.65	0.35	0.92	0.09	0.06	0.03	0.04	0.06	0.07	-0.39
LTGOVB	0.29	0.71	0.05	0.42	0.30	0.09	0.19	0.08	0.21	-0.67
HYC	0.50	0.50	0.03	0.05	0.02	0.22	0.43	0.10	0.18	-0.42
INTGOVB	0.40	0.60	0.03	0.64	0.23	0.48	0.14	0.24	0.50	-1.73
NAREIT	0.48	0.52	0.36	0.25	0.05	0.05	0.08	0.05	0.04	-0.07
NAREQ	0.53	0.47	0.26	0.19	0.09	0.03	0.01	0.03	0.03	0.18
NARMTG	0.53	0.47	0.19	0.19	0.13	0.02	0.27	0.02	0.02	0.00
NARHYB	0.43	0.57	0.30	0.19	0.06	0.04	0.17	0.02	0.04	-0.01
DJWILREIT	0.47	0.53	0.47	0.21	0.07	0.05	0.01	0.05	0.04	-0.08

Figure 1

First-Stage Fama-MacBeth Loadings on Macroeconomic Risk Factors: VW Market Portfolio

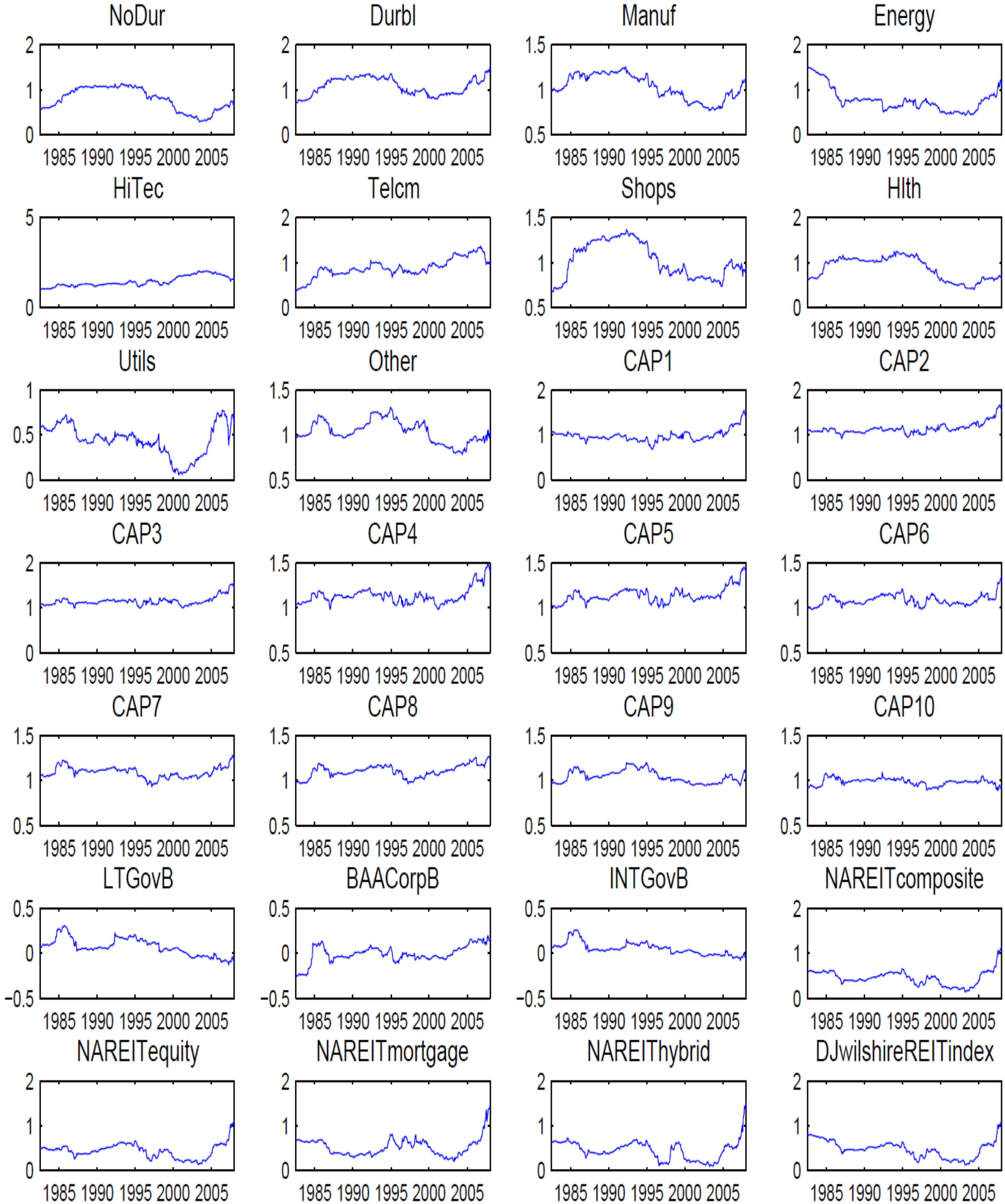

Figure 2

First-Stage Fama-MacBeth Loadings on Macroeconomic Risk Factors: Term Premium

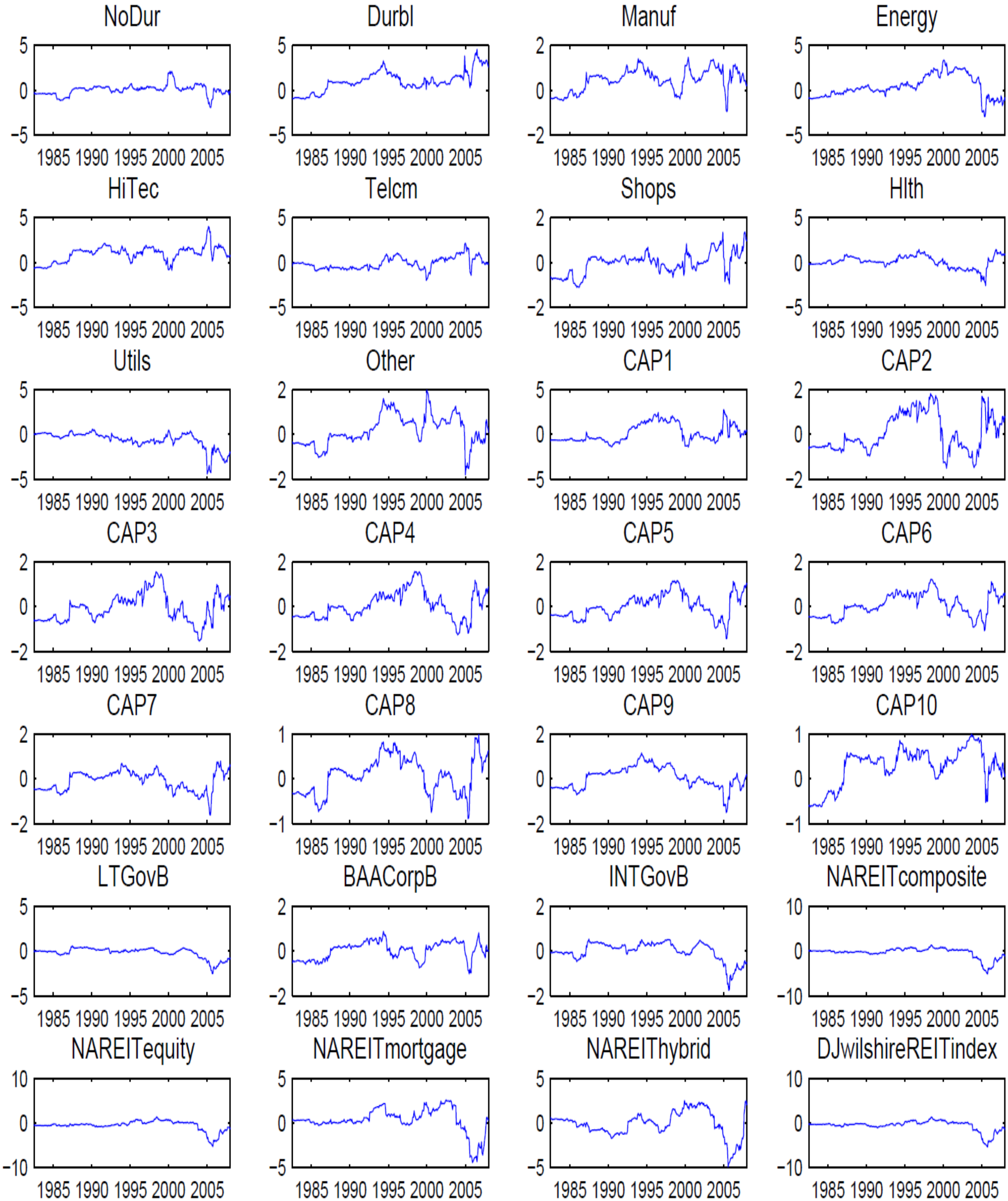

Figure 3

Bayesian Loadings on Macroeconomic Risk Factors: VW Market Portfolio

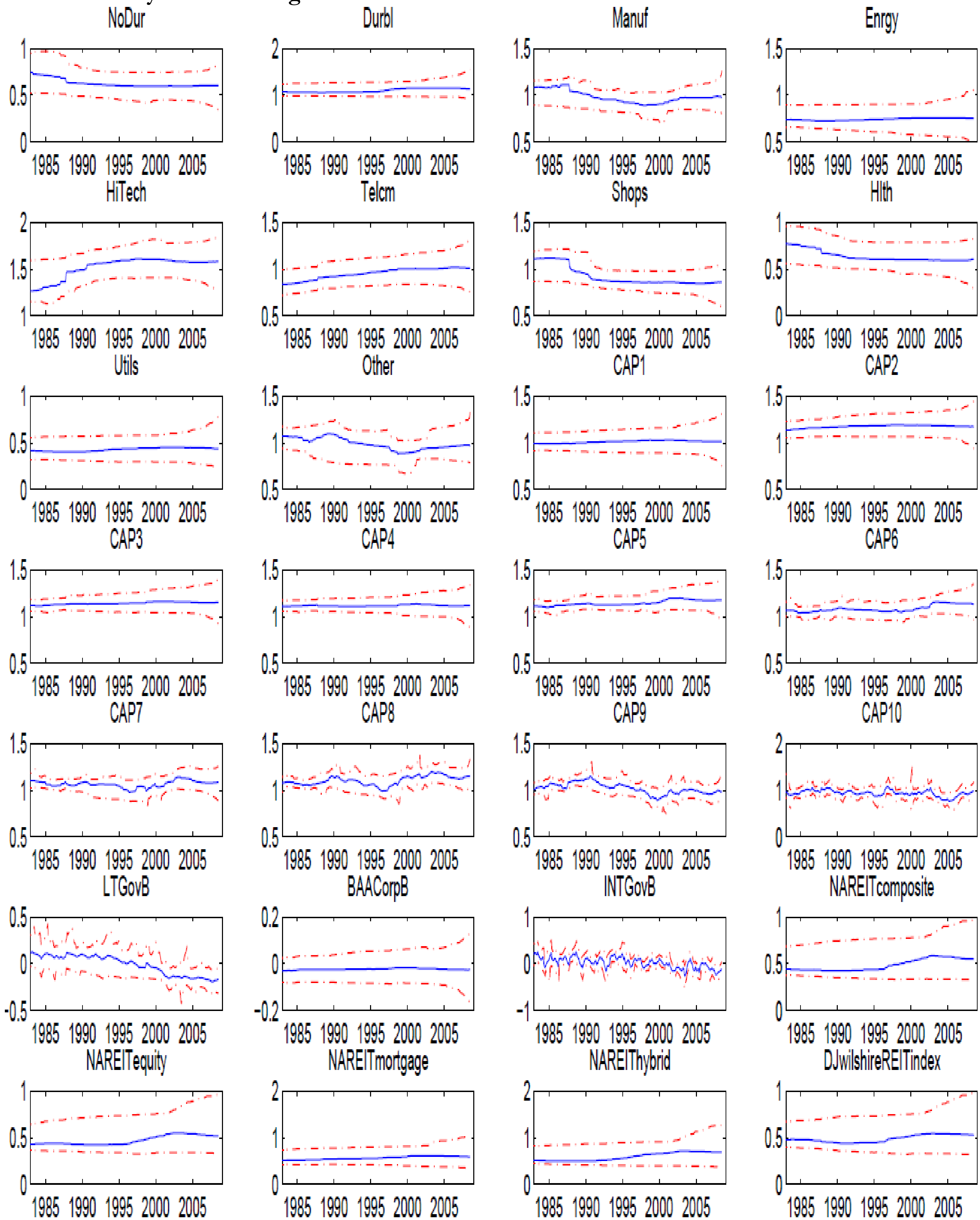

Figure 4

Bayesian Loadings on Macroeconomic Risk Factors: Term Premium

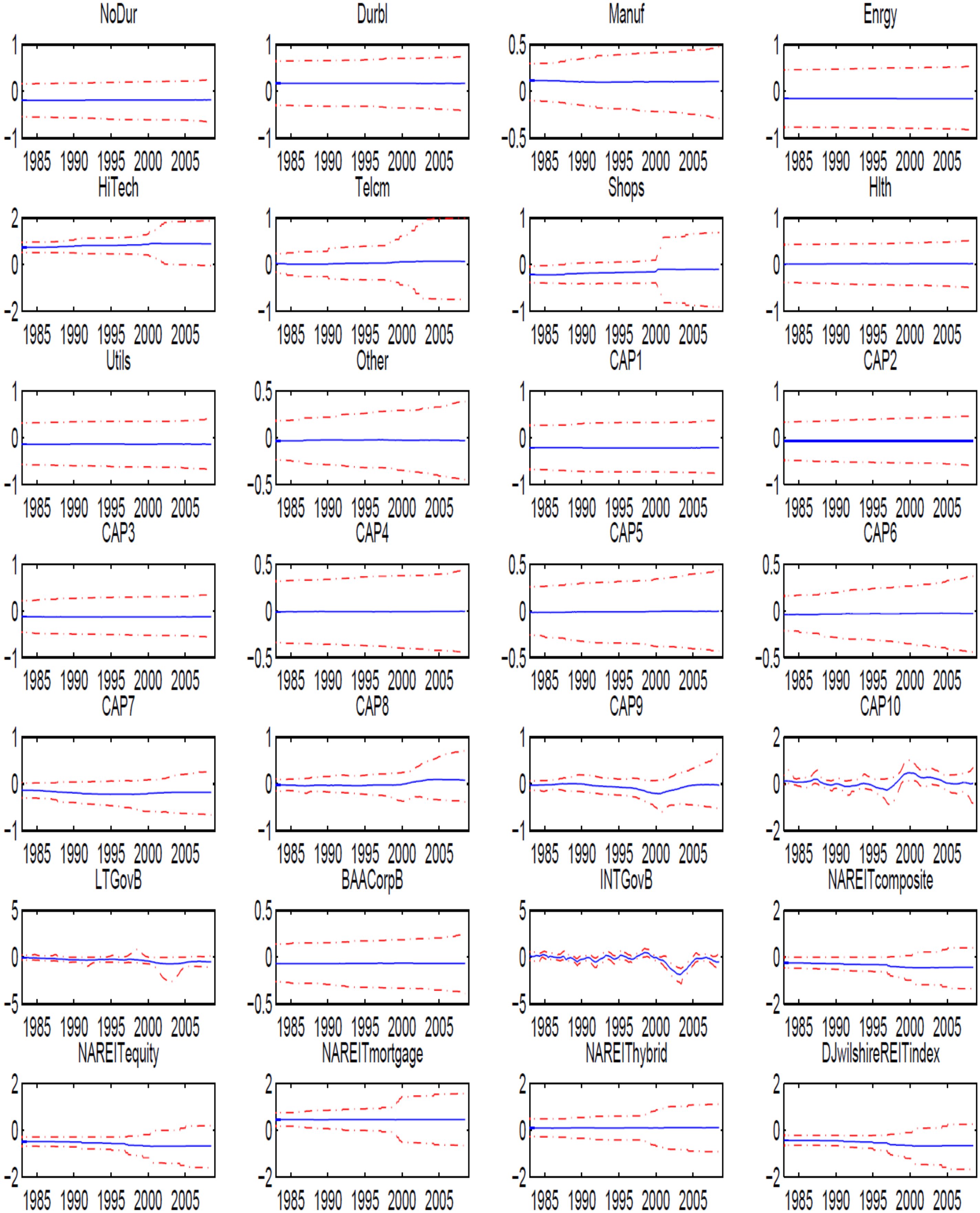

Figure 5

Bayesian Loadings on Macroeconomic Risk Factors: Default Premium

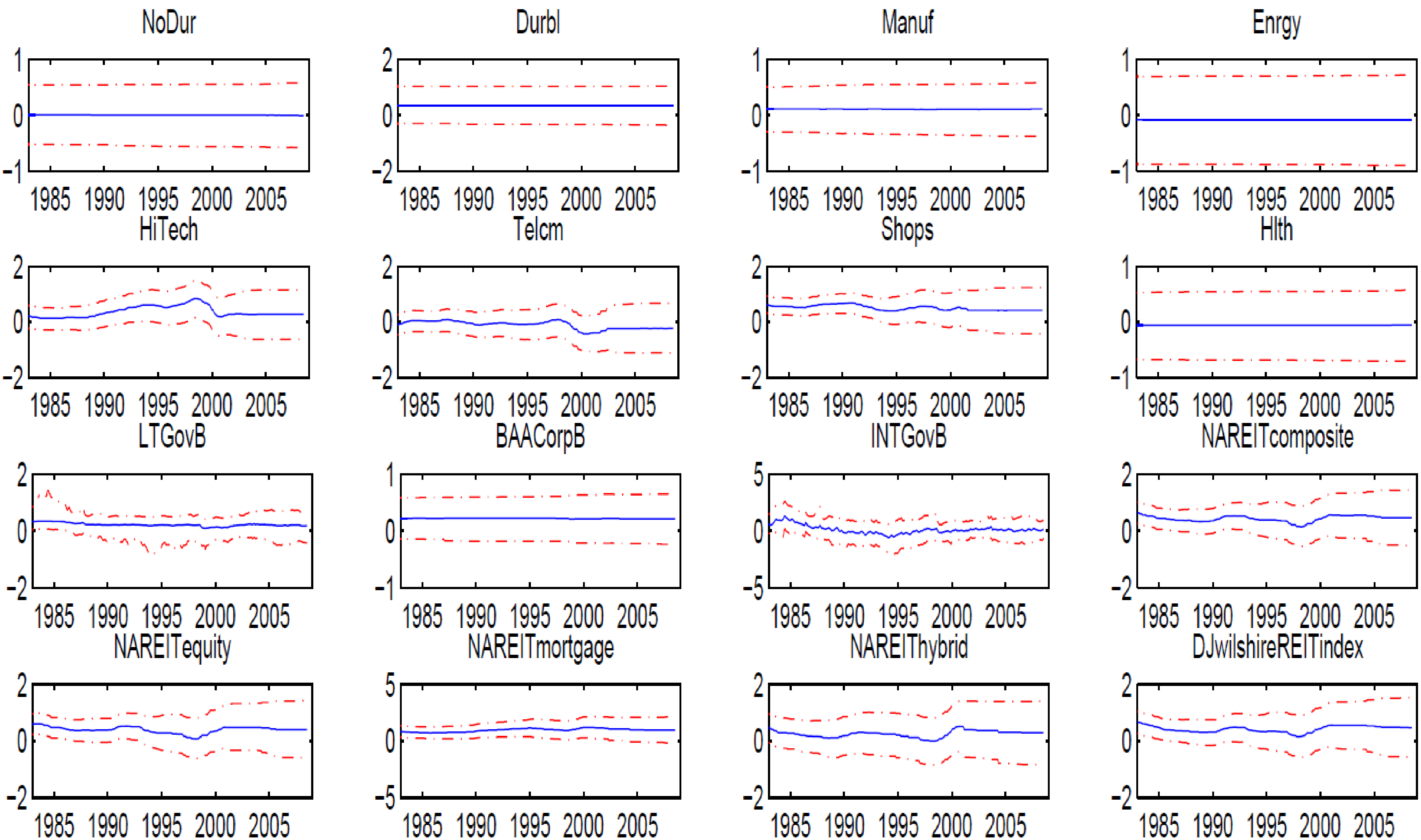

Figure 6

Bayesian Loadings on Macroeconomic Risk Factors: Real 1-month T-Bill Rate

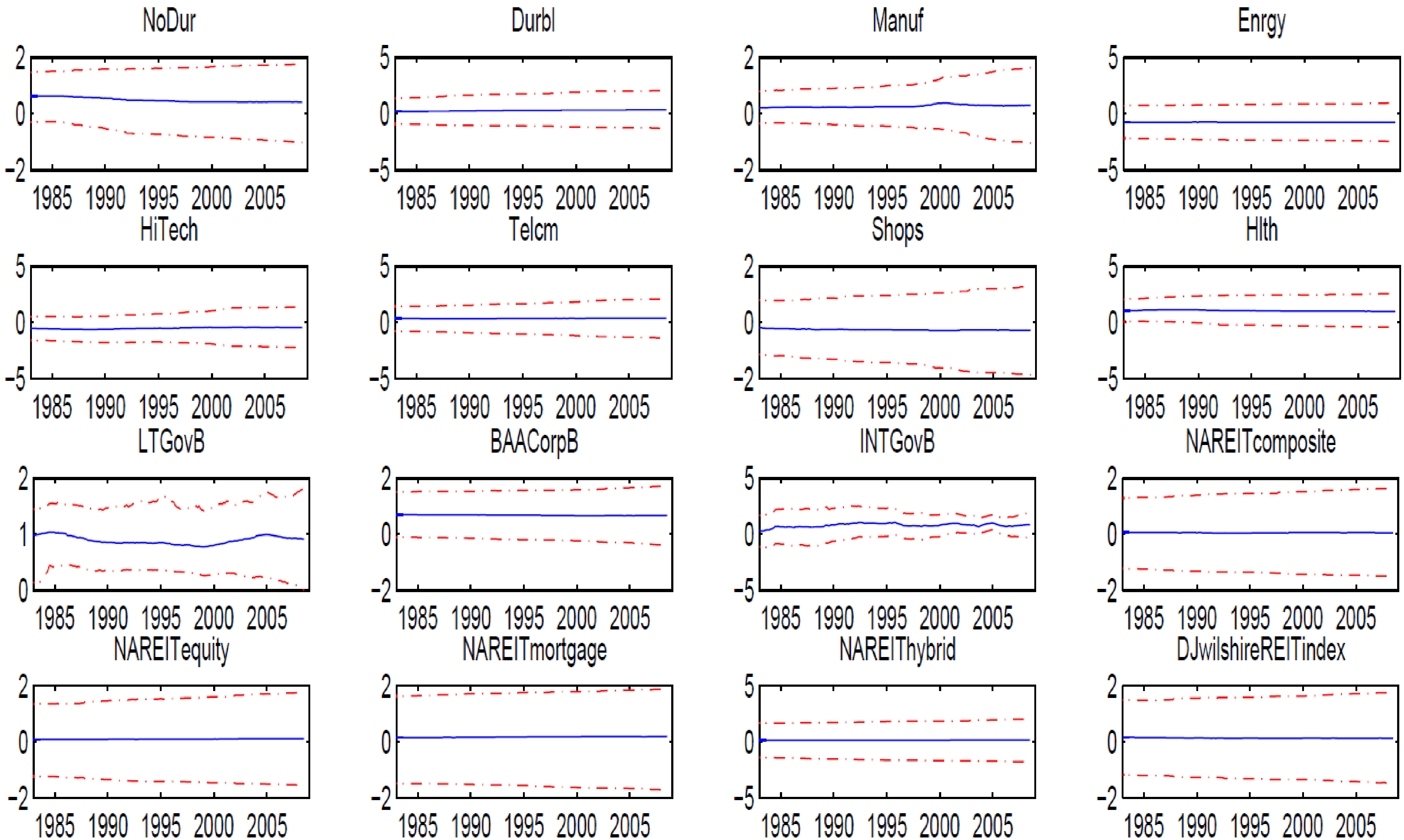

Figure 7

Bayesian Loadings on Macroeconomic Risk Factors: Industrial Production Growth Rate

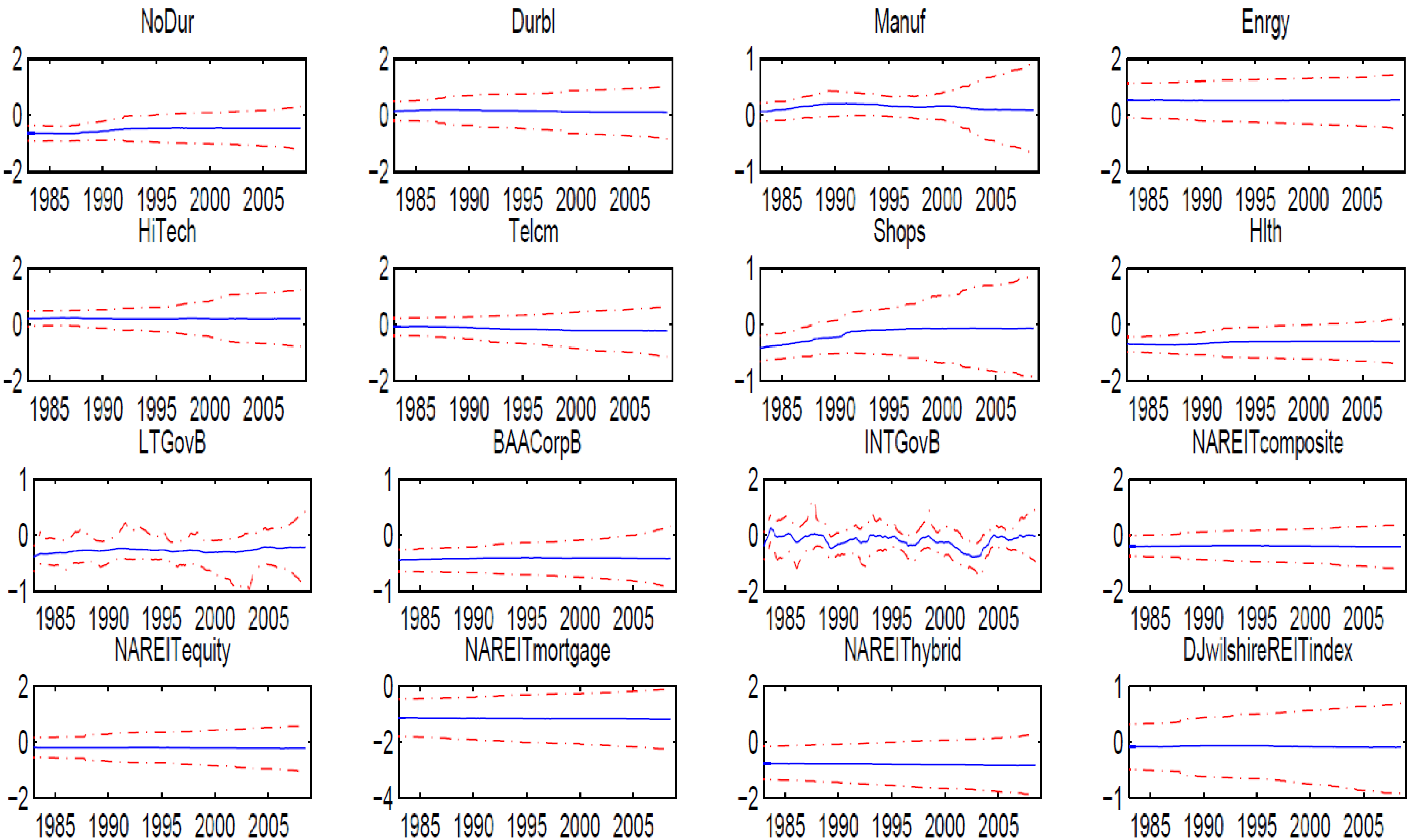

Figure 8
Bayesian Loadings on Macroeconomic Risk Factors: Changes in the Unemployment Rate

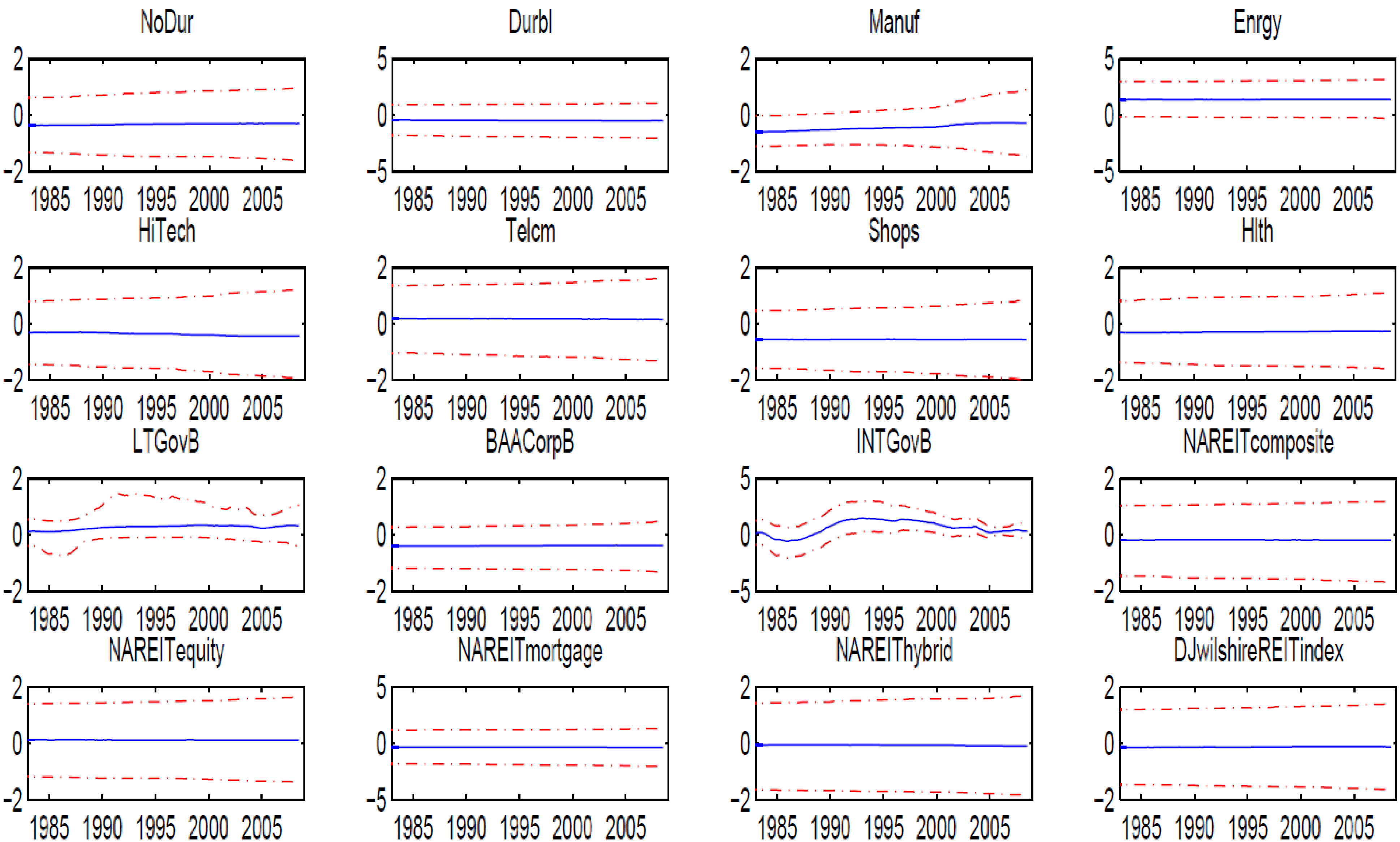

Figure 9

Bayesian Loadings on Macroeconomic Risk Factors: Real Personal Consumption Growth Rate

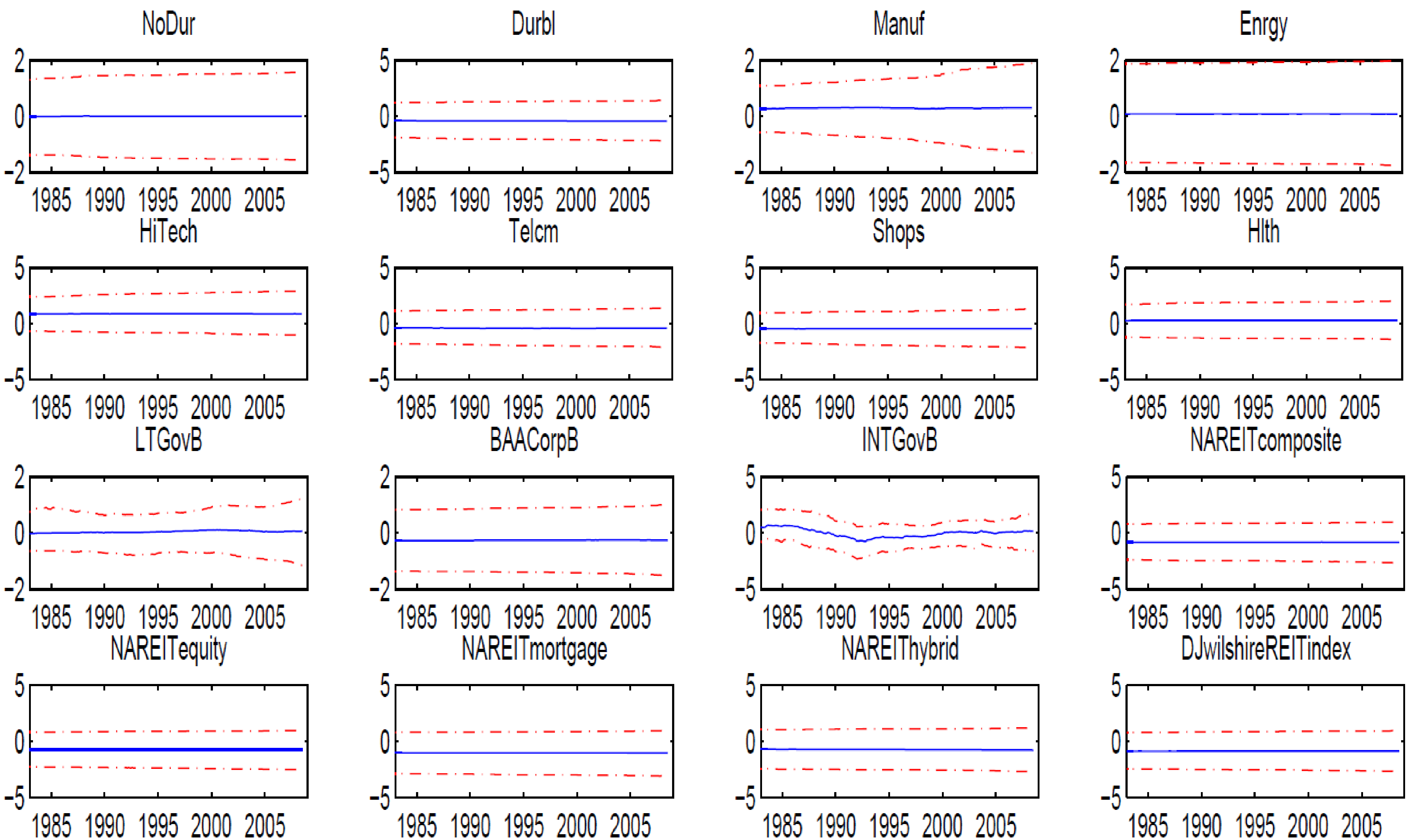

Figure 10

First-Stage Fama-MacBeth Rolling-Window Estimates of the Jensen's Alpha

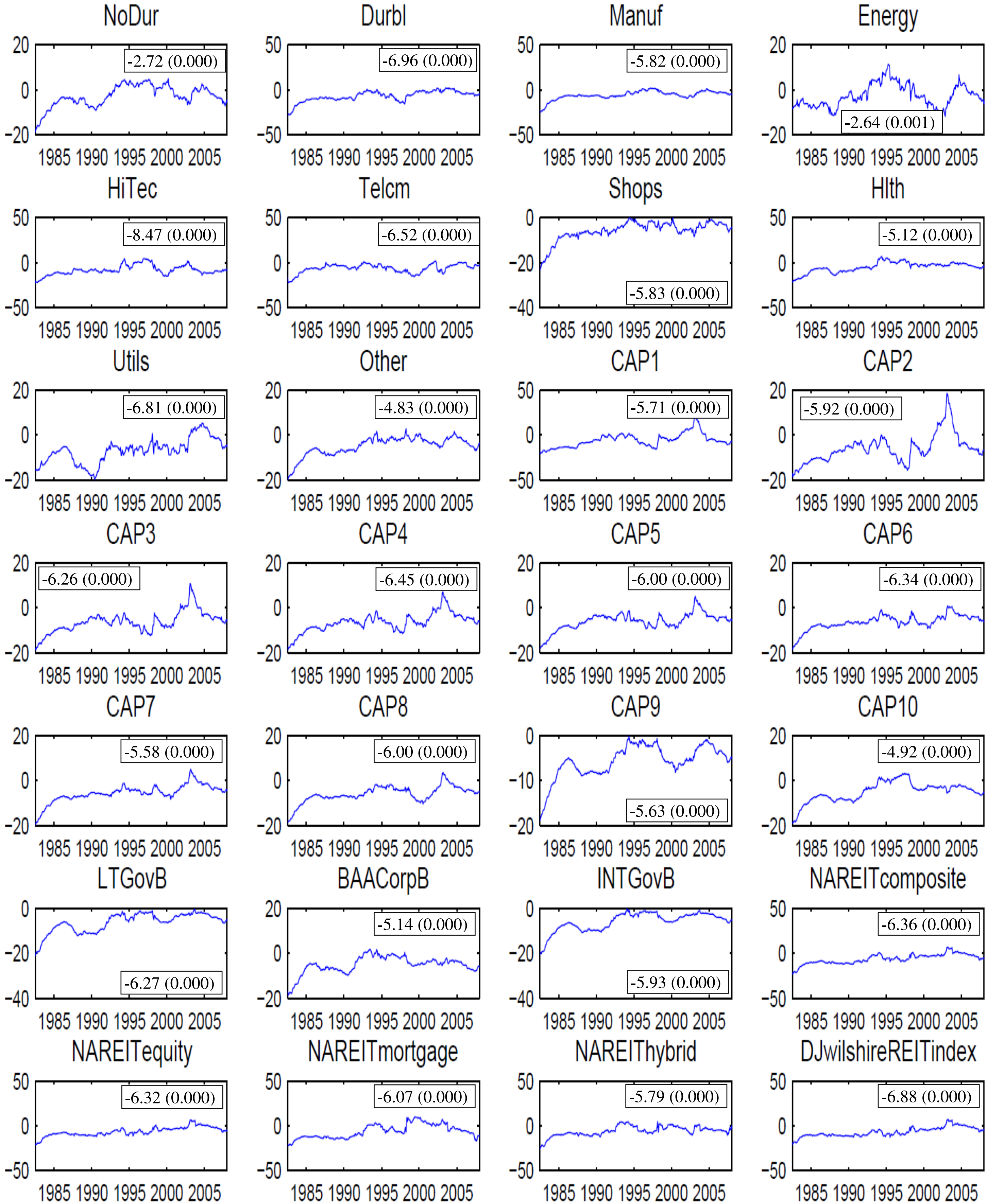

Figure 11

Bayesian Posterior Medians of the Jensen's Alpha

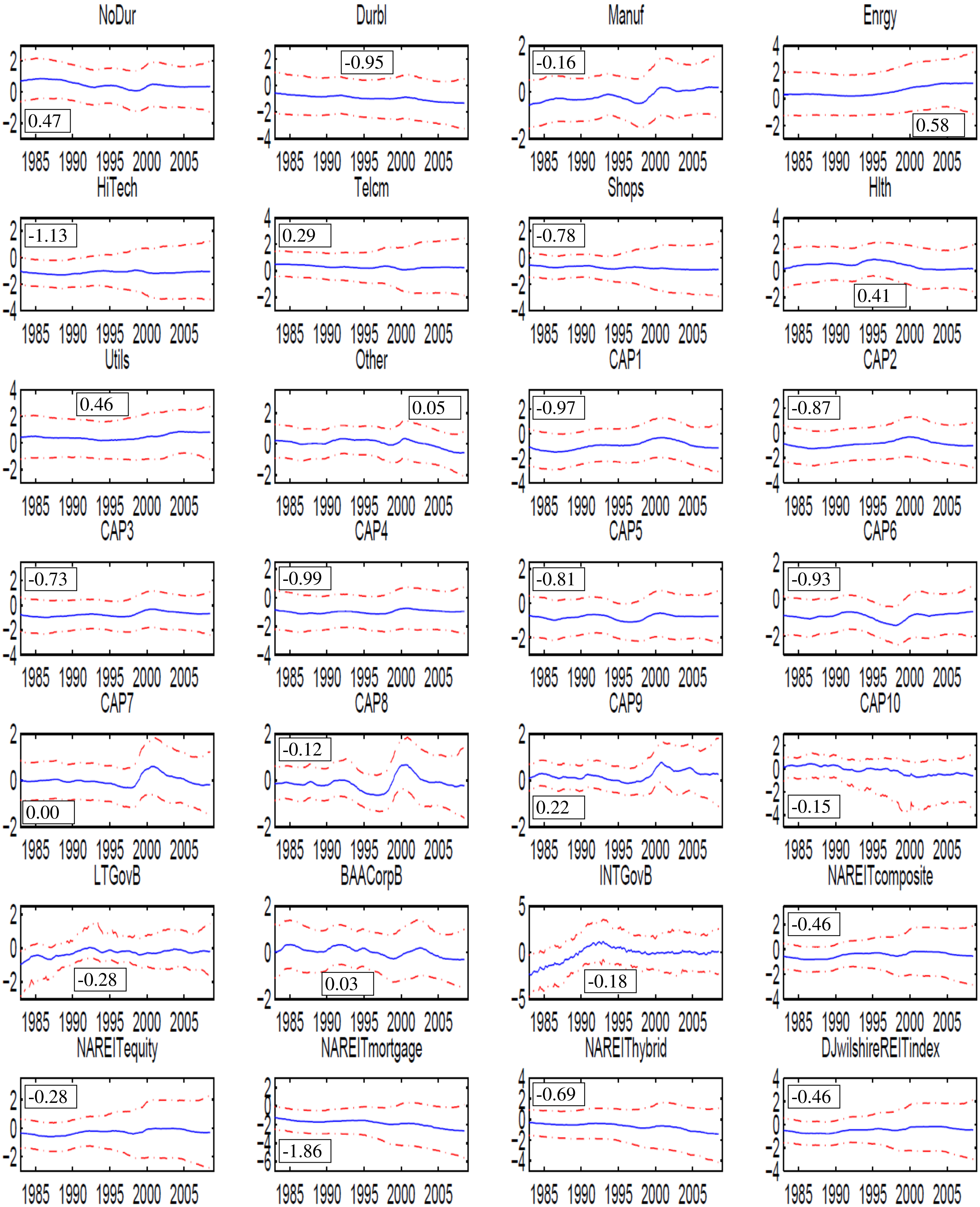

Figure 12

Fama-MacBeth Rolling-Window Estimates of Residual (Idiosyncratic) Variance

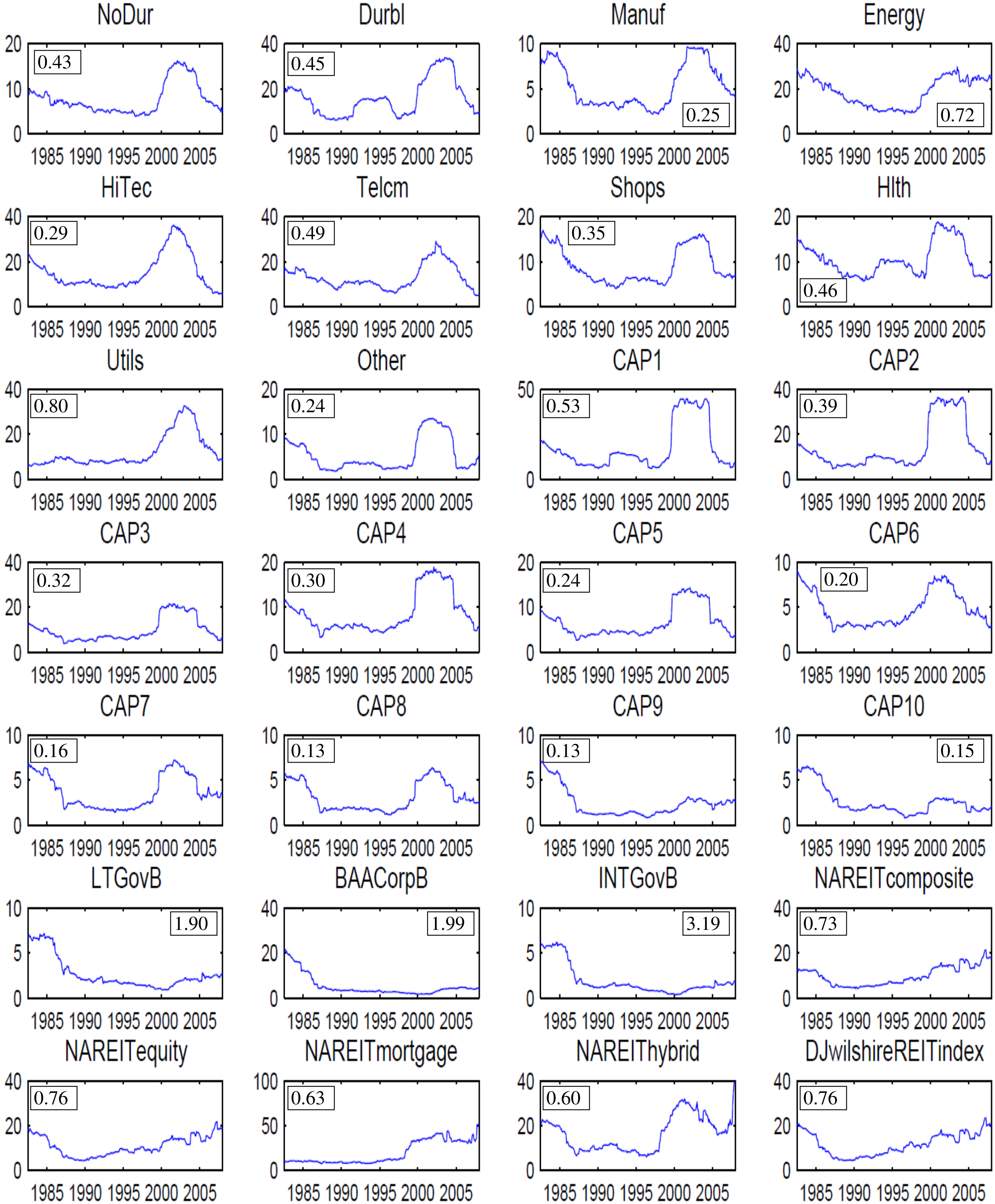

Figure 13
Bayesian Posterior Medians of Residual (Idiosyncratic) Variance

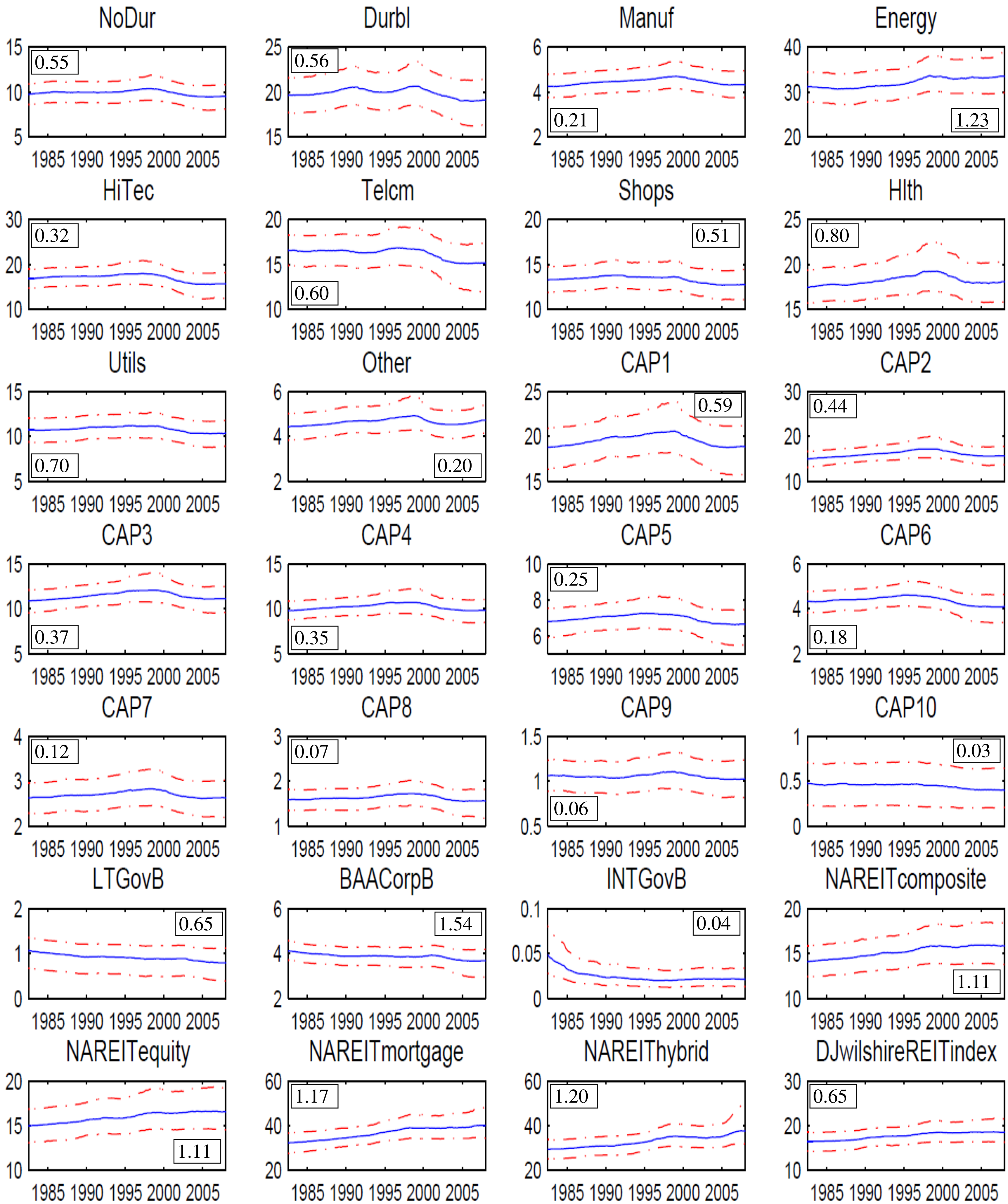

Figure 14
Fama-MacBeth Rolling-Window Cross-Sectional Estimates of Risk Premia

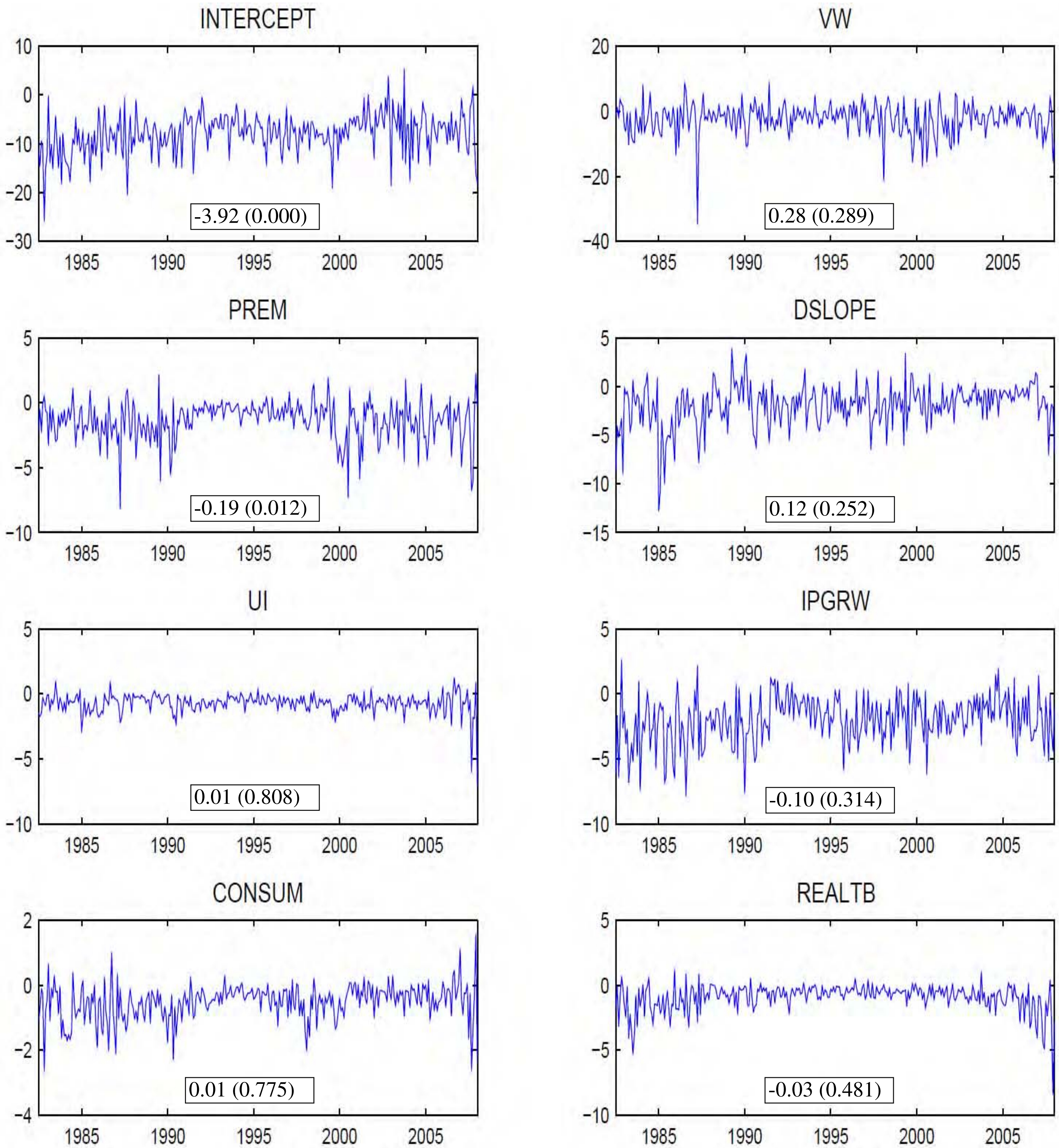

Figure 15
Bayesian Cross-Sectional Posterior Median Estimates of Risk Premia