DECLARATORIA SULLA TESI DI DOTTORATO

Da inserire come prima pagina della tesi

Il/la sottoscritto/a

COGNOME	TORTORA
NOME	ANDREA DONATO
Matr.	1094754

Titolo della tesi:

BAYESIAN MODELLING OF STRUCTURAL INSTABILITY: APPLICATIONS TO ECONOMICS AND FINANCE

Dottorato di ricerca in	ECONOMICS
Ciclo	XXI
Tutor del dottorando	PROF. CARLO AMBROGIO FAVERO
Anno di discussione	2010

DICHIARA

sotto la sua responsabilità di essere a conoscenza:

1) che, ai sensi del D.P.R. 28.12.2000, N. 445, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici previsti dalla presente declaratoria e da quella sull'embargo;

2) che l'Università ha l'obbligo, ai sensi dell'art. 6, comma 11, del Decreto Ministeriale 30 aprile 1999 prot. n. 224/1999, di curare il deposito di copia della tesi finale presso le Biblioteche Nazionali Centrali di Roma e Firenze, dove sarà consentita la consultabilità, fatto salvo l'eventuale embargo legato alla necessità di tutelare i diritti di enti esterni terzi e di sfruttamento industriale/commerciale dei contenuti della tesi;

3) che il Servizio Biblioteca Bocconi archivierà la tesi nel proprio Archivio istituzionale ad Accesso Aperto e che consentirà unicamente la consultabilità on-line del testo completo (fatto salvo l'eventuale embargo);

4) che per l'archiviazione presso la Biblioteca Bocconi, l'Università richiede che la tesi sia consegnata dal dottorando alla Società NORMADEC (operante in nome e per conto dell'Università) tramite procedura on-line con contenuto non modificabile e che la Società Normadec indicherà in ogni piè di pagina le seguenti informazioni:

- tesi di dottorato BAYESIAN MODELLING OF STRUCTURAL INSTABILITY: APPLICATIONS TO ECONOMICS AND FINANCE;

- di TORTORA ANDREA DONATO;
- discussa presso l'Università commerciale Luigi Bocconi – Milano nell'anno 2010;
- La tesi è tutelata dalla normativa sul diritto d'autore (legge 22 aprile 1941, n.633 e successive integrazioni e modifiche). Sono comunque fatti salvi i diritti dell'Università Commerciale Luigi Bocconi di riproduzione per scopi di ricerca e didattici, con citazione della fonte;

5) che la copia della tesi depositata presso la NORMADEC tramite procedura on-line è del tutto identica a quelle consegnate/inviate ai Commissari e a qualsiasi altra copia depositata negli Uffici dell'Ateneo in forma cartacea o digitale e che di conseguenza va esclusa qualsiasi responsabilità dell'Ateneo stesso per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;

6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi (legge 22 aprile 1941, n.633 e successive integrazioni e modifiche), ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto l'Università è in ogni caso esente da responsabilità di qualsivoglia natura, civile, amministrativa o penale e sarà dal sottoscritto tenuta indenne da qualsiasi richiesta o rivendicazione da parte di terzi;

7) che la tesi di dottorato non è il risultato di attività rientranti nella normativa sulla proprietà industriale, non è stata prodotta nell'ambito di progetti finanziati da soggetti pubblici o privati con vincoli alla divulgazione dei risultati; non è oggetto di eventuali registrazioni di tipo brevettale o di tutela, e quindi non è soggetta a embargo;

Data 01/02/2010

F.to ANDREA DONATO TORTORA

Bayesian Modelling of Structural Instability: Applications to Economics and Finance

Dissertation in partial fulfillment of the requirements for the academic degree of Doctor of Philosophy in Economics (XXI cycle).

Andrea Donato Tortora (1094754)
Università Commerciale Luigi Bocconi, Milan

Thesis Committee :

Carlo A. Favero, IGIER and Bocconi University

Massimo Guidolin, Manchester Business School and Federal Reserve Bank of St. Louis

Gianni Amisano, DG Research, European Central Bank and University of Brescia

Contents

Acknowledgements

I would like to express my gratitude to the many people that have contributed to the development of this dissertation.

First of all, I am particularly grateful to my main supervisor Prof. Carlo Favero for his guidance in shaping my ideas, his support and suggestions on several important issues throughout these years. Secondly, I thank my external supervisors. Prof. Massimo Guidolin has given me invaluable help and precious advice both during my stay at the Manchester Business School and later when back to Milan, being more than a supervisor. Prof. Gianni Amisano has initiated me to the Bayesian paradigm with extreme simplicity and has always been available to listen to my questions and to answer them.

Furthermore, a special thank goes to Dr. Francesco Ravazzolo, a coauthor of one of the essays in this dissertation, with whom I have had the chance of fruitfully exchange ideas and views. An essential ingredient for the development of a research agenda is the research environment in which the student happens to be. In this respect, I have had the opportunity to benefit from useful discussions with Bocconi faculty members and PhD colleagues. I am grateful as well to Bocconi University for the financial support I have been granted and the Manchester Accounting and Finance Group (MAFG) at Manchester Business School (MBS) for their kind hospitality during my research stay.

Finally, I thank my parents and my brother Giangaetano for their constant, extraordinary and unconditional support over these years, Elena and my friends in Milan and Giovinazzo, my hometown, for the love and the understanding they have always been willing to offer. Any mistake is obviously mine.

Introduction

This dissertation focuses on a prominent feature of many economic and financial time series, namely the presence of structural instability. Researchers typically use reduced form models whose parameters may shifts in response to changes in monetary policy rules, tax laws, regulation and so on. As a matter of fact, both inference and forecasting are severely undermined by the issue at hand. Hence, identification of the sources and nature of such time variation is key to sound economic decision making.

To this end, many models have been developed in the literature: for example, parameters may be postulated to undergo discrete or recurrent breaks and the latter may in turn be transitory or permanent. Among them, the time-varying parameters (TVP) model that allows for smooth time variation has become over the years the benchmark for its simplicity. Nonetheless, its appeal strongly diminishes as the dimensionality of multivariate systems increases. This is particularly true in forecasting applications in which very often the sample size under consideration is limited so as to render the estimation very unprecise relative to simpler models.

The present dissertation embraces a very flexible approach, referred to as "the mixture innovation approach", that does not impose *a priori* the number and the timing of structural breaks. Rather it encompasses both the fixed parameter model and the TVP model as special cases. A latent indicator, indeed, plays the crucial role to determine at each point in time whether a shift in parameters has taken place. Popularized by Giordani, Kohn and Van Dijk (2007), Ravazzolo, Paap, Van Dijk and Franses (2007) and Giordani and Kohn (2008), this new approach is garnering increasing attention. The practical estimation is carried out by means of Bayesian methods that allow to circumvent typical problems with maximum likelihood estimation and moreover take naturally into account another crucial aspect such as parametr uncertainty.

The first essay incorporates both model uncertainty and structural instability to predict exchange rate movements. These are well known to be hardly predictable and the literature has long sought a sound alternative to a no-change model. Notwithstanding, fluctuations in the foreign exchange markets are to some degree related to movements in fundamentals but these relationships are rather unstable and faint. The paper applies a method that endogenously choose among all the possible combinations of candidate predictors as in Wright (2008) and also let them exert a variable impact onto the exchange rate change by means of a "mixture innovation approach". Results for the US dollar *vis-a-vis* the Canadian dollar and the Japanese yen at a quarterly frequency show that there exists a relevant amount of model uncertainty and weighting predictions from all the models fares well compared to the benchmark for the out of sample period of 1990-2008. However, the presence of structural instability needs to be carefully modelled: the best results are obtained by keeping parameters constant or limiting the amount of breaks. The specific approach used in the paper has the merit of capturing rare but important changes and hence be less prone to overfitting concerns than time-varying parameters (TVP) models.

The second essay employs the "mixture innovation approach" for inferential purposes. In particular, the asset pricing properties and dynamics of publicly traded real estate are analyzed and compared to general asset classes like stocks and bonds. Real estate investment trusts (REITs) have been introduced in the United States in 1960 in order to make investment into large scale, income producing real estate accessible to small investors and have since been considered as an additional investment opportunity. A number of legislative interventions and the recent boom in capitalization have changed the properties of this asset class. Following the two-stage procedure *à* la Fama and MacBeth (1973), a multifactor pricing model is used to estimate, firstly, time-varying factor sensitivities and, secondly, risk premia in a specification in which also the conditional volatility is time-varying. The second step is performed in a pure Bayesian way, that is exploiting the entire posterior distribution of the estimated sensitivities as proposed by Ouysse and Kohn (2009) in order to account for parameter uncertainty. Results do not support the view of a sudden change in public real estate pricing characteristics around the early nineties. Yet, idiosyncratic risk shows a marked upward trend with a decreasing rate just in the last years while the quantity of priced market risk has risen since 2000. The time-varying multi-factor pricing model turns out to be rejected for the asset menu under consideration.

The last essay focuses on the predictive relationship that links the term spread, measured as the difference between the yields on long term (10-year) and short term (3-month) government bonds, to future output growth. The former is historically considered to be a reliable harbinger of future recessions but, on the other hand, its performance has been documented to have decreased in the last twenty years. The aim is to include recent observations, precisely data until 2009:II, to the overall picture. To capture these features, a plain TVP model that allows for changes in regression parameters, conditional variances and correlations as in most macroeconomic studies is used. The analysis is not restricted to a bivariate model but computes the marginal predictive content of the yield spread once inflation and the short term rate are added as well. This measure of fit at one and two years ahead is inevitably computed in-sample as the dimensionality of the system makes a recursive approach unfeasible. The evidence confirms the findings in the previous literature and also suggests that such predictive power has not moved much lately, apart from a short-lived hike around the time of the recession in 2001. Furthermore, it appears more important the contribution of inflation and the short rate.

References

[1] Fama E. and MacBeth J., "Risk, Return and Equilibrium: Empirical Tests", *Journal of Political Economy*, 1973, 81, 607636.

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

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

[4] Ouysse R. and Kohn R., "Bayesian Variable Selection and Estimation of Risk Premiums in the APT model", *Discussion Papers*, School of Economics, The University of New South Wales, 2009.

[5] 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.), "Frontiers of Economics and Globalization" Volume 3, *Forecasting in the Presence of Structural Breaks and Uncertainty*, Emerald Publishing Ltd. & Elsevier Press, 2007, pp. 561-594.

[6] Wright J. H., "Bayesian Model Averaging and Exchange Rate Forecasts", *Journal of Econometrics*, 2008, N. 146, 329-341.

Exchange Rate Forecasting: Bayesian Model Averaging and Structural Instability

Abstract

This paper addresses the topic of exchange rate forecasting by using a Bayesian Model Averaging approach which explicitly accounts for both model and parameter uncertainty. Wright (2008) is the first exploring this route and provides encouraging results. The novelty here consists of adding one more ingredient, parameter instability, by means of a mixture innovation approach [see Ravazzolo, Paap, van Dijk and Franses (2007) and Giordani and Kohn (2008)]. Besides analyzing shortly both model uncertainty and parameter instability in the full sample, their relative contribution is evaluated *versus* the benchmark driftless random walk model for one-quarter-ahead predictions of two exchange rates. The statistical criteria used for comparisons explore several dimensions, not only the mean of the forecast distribution. The proposed averaging approach works well for the US dollar - Japanese yen exchange rate, while for the US Dollar - Canadian Dollar case the performance appears to depend on the specific time window under consideration. Overall, the best results are obtained by keeping parameters constant or limiting the amount of breaks: hence, the mixture innovation approach appears superior to the time-varying parameter (TVP) model framework as a forecasting tool in the present application.

1 Introduction

The seminal Meese and Rogoff's paper (1983) has posed a serious challenge for any macroeconomic model of exchange rate determination by decreeing the primacy of the random walk specification. This finding, termed as the "exchange rate disconnect puzzle", plainly implies that returns are unpredictable given current information. The countless attempts to fix it have been mostly unsatisfactory: indeed whenever different results emerge they just hold for either a specific sample or a particular currency.

Among the routes followed in the literature, forecast combination techniques have garnered increasing attention recently, examples being Altavilla and de Grauwe (2010) and Della Corte *et al.* (2009). Bates and Granger (1969) have been among the first ones suggesting the high potential benefits from forecast combination. In general, as Guidolin and Na (2007) recognize, it is well accepted in the literature that, in case of structural instability, single nonlinear models are not likely to perform always accurately while forecast combination can provide a hedge against nonstationarity.[1] Bayesian Model Averaging offers a rigorous statistical foundation to such a practice as each forecast is weighed by the posterior probability of the respective model.

In the present context Wright (2008) (henceforth, Wright) finds that forecasts from Bayesian Model Averaging sometimes perform quite better than the random walk while they never do much worse, even though it turns out that the two competing forecasts are quite close. His analysis is here extended by introducing time-variation in the relationship between dependent and independent variables. Even though long considered as a possible solution to the Meese-Rogoff puzzle, the contribution offered by this kind of instability has usually been neglected, the few exceptions being Schinasi and Swamy (1989) and Rossi (2006). To this purpose, a mixture innovation approach as in Ravazzolo *et al.* (2007) and Giordani and Kohn (2008) is used: it is very flexible in that parameters are allowed to change at each time but, crucially, they are not forced to do so. This draft deals with one-quarter ahead predictions, exactly the horizon for which traditional macromodel-based forecasts are mostly disappointing with respect to a

[1] There is a vast ongoing literature investigating forecast combination techniques and comparing relative merits: a recent but interesting example is Stock and Watson (2005). The main alternative to Bayesian Model Averaging in the frequentist approach is "bagging" as developed by Inoue and Kilian (2008).

random walk.[2] The focus is mainly, but not solely, on point predictions. In fact I also look at other important aspects like distribution forecasts, even though these are just briefly touched upon and left for a more detailed future analysis.

Over the main evaluation sample (1990-2008), results are promising for the US dollar-Japanese yen but not for the US dollar-Canadian dollar. Interestingly, focusing on the same out-of-sample window as in Wright (1990-2005) the latter conclusion is partially reversed as models with modest and rare time variation in parameters fare better than the random walk.

Hence the message is that Bayesian Model Averaging may help if carefully designed, in particular as to how parameter instability is modelled. In fact, the best performing strategy turns out to be averaging over the space of models with constant (in other terms, similarly to Wright), or slightly and occasionally moving parameters. When the same models are given too much flexibility (TVP models) the quality of the resulting predictions easily deteriorates. This result is not new in the literature, in particular for models with a random-walk-type variation: in a more general and systematic investigation, Stock and Watson (1996) find widespread instability in macroeconomic variables but recognize however that TVP models often fail to exploit it as they are outperformed by constant parameter models. In the case of floating exchange rates, the amount of information available at quarterly frequencies hardly guarantees precise estimates of parameters that are subject to changes over time, let alone model misspecification. Especially when high-dimensional models receive some non-negligible prior probability, like in the proposed approach. In short, the mixture innovation approach appears a preferable alternative to smooth time variation in similar environments characterized by instability.

Moreover, the well estabilished result in the forecasting literature that parameter shrinkage works is confirmed here as well. Finally, moving from statistical to more profit-based metrics like the Hit Rate, the overall picture becomes clearly more favourable.

The remainder of the paper is organized as follows. Section 2 briefly resumes the state of the art in the empirical exchange rate literature, Section 3 introduces the model structure, discusses

[2]This common view can anyway be questioned in light of recent evidence such as Molodtsova and Papell (2009) (see next section). On the contrary, the good performance in the long horizon is a well estabilished result in the literature.

the prior choices and illustrates the algorithm used for posterior simulation. Section 4 present data and results while Section 5 concludes.

2 Empirical Exchange Rate Literature[3]

One of the most controversial issues in international finance concerns the faint relationship between economic fundamentals and exchange rates for advanced countries, also known as "the exchange rate disconnect puzzle". A vast set of theoretical models, ranging from the monetary approach to the portfolio balance approach as well as the recent elaborate "new open macro" frameworks, highlights the relevance of macroeconomic fundamentals such as money or output differentials or current account balances in determining the exchange rate level. Yet, none of them has managed to pass simple tests like out-of-sample forecasting performance. Except for some specific circumstances, the random walk tends to outperform monetary or "fundamental" models whatsoever in short-horizon regressions while loosing often in the long-horizon, as shown by Mark (1995).[4]

Nonetheless, renewed interest has been spurred by recent contributions that have partially rehabilitated traditional models and criticized the mainstream literature in many directions. For instance, Engel, Mark and West (2007) (hereafter, EMW) provide both theoretical and empirical support to traditional macro models despite their well-known out-of-sample failure in outperforming a random walk. The main argument relies on the fact that short-run exchange rate changes are driven mostly by changes in expectations about future fundamentals and according to these models, under weak conditions, exchange rates behave similarly to a random walk. As a consequence, different tools other than out of sample tests have to be used for evaluation purposes.[5]

[3]The following section is intended to outline the recent contributions of an ever prolific literature. Most of them are anyway only of indirect interest as they aim at explaining empirical failures of specific theoretical models, whereas the focus in this paper is broader and more practical.

[4]However, the implementation in such exceptions raises some doubts about their econometric soundness and the reliability of their results. A thorough analysis is in Neely and Sarno (2002): please refer to details therein.

[5]EMW list several criteria that give support to traditional rational expectation macro models but also highlight a relatively new evidence about both the usefulness of Taylor rule based predictions for exchange rates and the presence of a Granger causality relationship with future fundamentals.

Clark and West (2006) instead point to the drawbacks with using the Diebold-Mariano (hereafter DM) statistics for inference in case of nested models. They show that the DM procedure is severely undersized and has also low power: even though the MSPE difference in population is zero, the in-sample difference is negative. The same authors adjust the statistic for this upward shift (henceforth CW statistic) and find properly sized tests by using asimptotically normal critical values. A bridge between EMW and Clark and West (2005) is Molodtsova and Papell (2009) who point to the effect of model selection and inference methodology on out-of-sample predictability. In particular, they use the CW statistic and find evidence of short term exchange rate predictability especially when using Taylor rule models. However, the predictive power dies completely away beyond a six month horizon.

Another explanation for the random walk outperformance refers to the presence of underlying structural changes, in the form of simple breaks, regime switches or nonlinearities in the relationship between exchange rates and fundamentals. Structural instability characterizes indeed most of macroeconomic and financial series and can easily produce conflicting results between in sample and out of sample tests.[6] Rossi (2006) investigates the implications of parameter instability for model selection between fundamental-based and no-change models of exchange rate determination: it emerges an unstable relationship between the exchange rate and the fundamentals and in some cases (Japanese Yen-US Dollar, for example) accounting for breaks delivers better forecasts than a simple random walk. This paper shares with Rossi (2006) the intuition that exploiting such features can be fruitful for forecasting but carries out a Bayesian rather than a classical analysis.

A related idea is originally put forward in Sarno and Valente (2008): based on survey data evidence of swings in expectations, they contend that traditional fundamentals do have some predictable content but the market attaches variable weights to them over time. Using *ex-post* data to select the best model at each time outperforms a random walk model for three out of five exchange rates. Hence the authors conclude that the kinds of failures reported in the literature

[6] In fact Meese and Rogoff (1988) consider in-sample tests unreliable and thus prefer out-of-sample procedures. In general, stable structural relationships are needed in order to rely on inference based on short time series, otherwise the out-of-sample forecasting performance will always be disappointing. For a recent investigation see also Clark and McCracken (2006).

need to be ascribed to weaknesses in typical model selection criteria.[7]

On the other hand Bacchetta, van Wincoop and Beutler (2009) challenge the role of parameter instability to explain the Meese-Rogoff puzzle. They use a fairly general reduced form model, very similar to the one laid out in the next section, and by both estimation and model calibration find that it is the small sample estimation bias that determines the puzzle.[8]

Finally, Rogoff and Stavrakeva (2008) temper any kind of optimism in recent studies by investigating the statistical significance of such macro-model forecasts. Several key aspects are pointed out. Firstly, unlike the Theil's U an the Diebold-Mariano statistic, it is not guaranteed that the new kind of tests (CW and ENC-NEW) pick the best model in the sense of minimum MSFE. In particular, in case forecasts are biased these newer out-of-sample tests simply impose the null that the DGP for exchange rates is a random walk. Moreover the authors test the robustness of positive results to different forecast windows. Critically, the best results are for some (Australia and Canada) of the "commodity currencies" and come from no structural factors. The main message can be resumed by shortly quoting a paragraph of their introduction:

> " The euphoria has been exaggerated by misinterpreation of some newer out-of-sample statistics for nested models, over-reliance on asymptotic out-of-sample tests statistics and failure to check for robustness to the time period sampled."

3 The approach: Bayesian Model Averaging and Structural Instability

3.1 The model

Model Averaging is a very useful tool for inference, prediction and policy analysis whenever there exists uncertainty about the true data generating process or theoretical model, this event

[7] The models selected according to the their strategy are clustered over time and none of them is the best one for long periods of time: this brings about shifts in the parameters that reconcile with evidence and explanations advanced in much of the recent literature.

[8] Anyway, the possibility that parameters have a unit root as in (5) is purposely ruled out: in this case, the puzzle would not exist at all according to their results. Furthermore, note that the objective of the present analysis is not to discover the true data generating process but to test the short horizon predictive power of macroeconomic variables.

being very likely to come along in economics. As put forward by Draper (1995), conditioning results on a specific model, as it is usually done, brings about biased results and too narrow standard errors. In a model averaging perspective, results are instead obtained by appropriately weighting the evidence from every model in the pool.

The idea of Bayesian Model Averaging (henceforth, BMA) traces back to Leamer (1978) even though the number of applications to econometrics is very small. If the researcher believes that a particular phenomenon or variable is driven by different factors or explainable in terms of various competing models, then it is preferable in many instances to explicitly account for this uncertainty instead of basing inference on a particular specification. Considering a set of models $M_1, ..., M_n$, it may be that the researcher does not know which one is the true model but she has prior beliefs about it, represented by $P(M_i)$. The relevant quantity becomes the posterior probability associated to each model:

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{\sum_{j=1}^{n} P(D|M_j)P(M_j)} \tag{1}$$

where

$$P(D|M_i) = \int P(D|\theta, M_i)P(\theta|M_i)d\theta \tag{2}$$

is the marginal likelihood of the $i - th$ model, $P(\theta|M_i)$ is the prior density of the parameter vector in that model and $P(D|\theta, M_i)$ is the likelihood. The resulting forecast density is

$$f^* = \sum_{j=1}^{n} P(M_i|D)f_i \tag{3}$$

where f_i is the forecast density from the $i - th$ model. Also, one can only look at point forecasts and thus weighting them by the respective posterior model probability. The elicitation of the parameter priors requires careful attention: in particular, improper priors are allowed only for those parameters common to all models for the marginal likelihood not to depend on arbitrary normalizing constants.

It is customary to allow for all possible combinations of likely explanatory variables, thus considering 2^m models in case of m regressors: when m is very large, for example more than 20, then the evaluation of each single models becomes unfeasible and an algorithm for searching over the model space is needed. This could be either deterministic or stochastic.[9]

The most recent applications relate to cross-country growth regressions [Doppelhofer, Miller and Sala-i-Martin (2000), Fernandez, Ley and Steel (2001), Cuaresma and Doppelhofer (2006)], stock market prediction [i.e. Avramov (2002), Cremers (2002) and Ravazzolo, Paap, van Dijk and Franses (2007) and Della Corte, Sarno and Tsiakas (2009)], and inflation prediction [Wright (2009), Jacobson and Karlsson (2004), Koop and Potter (2006) and Eklund and Karlsson (2007)], to name a few.

Ravazzolo *et al.* (2007) allow for both model uncertainty and parameter instability, features that have always been analyzed separately. The latter is introduced by means of a flexible approach that lets (but not binds) the parameters change at each point in time. Fixed parameter models and TVP models are both encompassed as special cases. Their approach has been followed here and the resulting regression model is the following:

$$\Delta e_t = \beta_{0t} + \sum_{j=1}^{m} s_j \beta_{jt} x_{jt-1} + \epsilon_t \tag{4}$$

where

$$\beta_{j,t} = \beta_{j,t-1} + k_{j,t}\eta_{j,t}, \qquad j = 0, ., m. \tag{5}$$

The latent binary random variables s_j and $k_{j,t}$ determine the inclusion of x_{jt} in the model and the presence of changes in $\beta_{j,t}$, respectively, while $\epsilon_t \sim N(0, \sigma^2)$ and $\eta_{j,t} \sim N(0, q_j^2)$. Each predictor variable x_j is included in the model with probability $Pr[s_j = 1] = \lambda_j$ and each model is denoted by the sequence $(s_1, ..., s_m)$. The regression parameter $\beta_{j,t}$ undergoes a change with

[9] The range of possible algorithms is very wide and so the interested reader is invited to read either Raftery *et al.*(1999) or Koop (2003) for a useful survey. A general reference for applying BMA to linear regressions is Raftery *et al.* (1997).

probability $Pr(k_{j,t} = 1) = \pi_j$ for $j = 0, ..., m$, which is independent of the dynamics up to that time period: if $k_{j,t} = 1$ the variation is then determined by $\eta_{j,t}$, otherwise it is null. Furthermore, the latent indicator is uncorrelated across variables which implies that the regression parameters are not restricted to change contemporaneously. The intercept is always included and is also allowed to change over time: in particular, regressors have to be demeaned in order for changes in the parameters not to influence the intercept.

3.2 Prior specification

A crucial and thoughtful step in any Bayesian estimation exercise is to elicit prior information. When entertaining a vast number of candidate models, a detailed specification of prior uncertainty is almost unfeasible and so a set of uninformative/automatic priors is advocated.

As for the model space, I follow the standard choice of a uniform prior which equally supports all the possible models:[10] the variable inclusion probabilities $\lambda_1, ..., \lambda_m$, where $Pr[s_j = 1] = \lambda_j$, are set equal to 0.5 for all $j = 1, .., m$. As a consequence, posterior model probabilities turn out to be proportional to marginal likelihoods. In the forecasting exercise I follow Wright in that I also experiment with $\lambda = 0.2$ which gives greater prior support to low-dimensional models.

For the measurement error variance, a natural choice is the independent conjugate prior like the inverted Gamma distribution. In particular, I impose it be centered close to the sample variance of the dependent variable and have a minimal amount of tightness (in this case $\nu = 2$).

A prior distribution for the β's is needed to initialize the filtering recursions: a normal distribution with mean zero is the natural candidate. As for its dispersion, I have tried several options that imply some moderate strenght for the proposed value.[11]

As for the peculiar elements of the approach used in this paper, the literature lacks many references unlike the abundant collection of works on time varying parameters. Among the exceptions is Koop, Leon-Gonzalez and Strachan (2009) (henceforth KLS) who adapt Primiceri

[10]Note that the uniform distribution do put structure on features like model size: the higher the inclusion probability of all variables, the bigger the expected model size. Here, $\lambda_j = 0.5\ \forall j$ implies an expected model size equal to $0.5 * m$, where m is the total number of predictive variables considered. Some might prefer to reward more parsimonious specifications.

[11]An important reason why not to leave the prior variance unconstrained is expressed in footnote 17. The reader interested in the actual hyperparameters used in the forecasting exercise can look at Tables 7 and 9.

(2005)'s settings, currently a benchmark in the TVP literature, to the specific framework at hand. The latter suggests using a mild informative prior about the break size that he centers on a specific fraction of the OLS coefficient variance (in his paper such a multiplicative factor equals 0.01^2) in order to rule out implausible regions of the parameter space. KLS, in turn, multiplies the same fraction by the inverse of the prior mean break-occurrence probability: in short, if breaks are for instance believed to occur on average every two quarters, the associated prior shock variance will be centered around two times the fraction suggested by Primiceri (2005). I do the same here but the fraction I consider is computed on the prior variance of the initial condition rather than the OLS slope variance.[12]

The probability of observing a structural break for the $j-th$ variable, described by π_j, is distributed as a Beta random variable: its hyperparameters, c_j and d_j, have to be set according to the subjective information about the duration mean between breaks. KLS impose the minimal tightness ($c_j = d_j = 1$) while Giordani and Kohn (2008) propose a prior that favors few breaks. The former is a more appealing choice to let the data speak in a framework in which the influence of prior information is already relevant, but it lends itself to likely problems of overfitting that could then bring about very poor forecasts. I have found that with such an option the algorithm becomes unstable on some occasions for the Canadian case while it works well for the Japanese yen. Then I have opted for a tighter prior still centered on the same value for break probabilities. Besides, for the analysis in section 5.2, I have mainly used a moderately informative prior implying a five percent break probability ($c_j = 5$ and $d_j = 95$). The reason for this is to push (but not force) the sampler to find at least a break already in the first prediction, which takes place at the last quarter of 1989.

A general description of prior distribution choices is provided in Table 1.

[12] I opt for a generic degree of dispersion that is not data-based. In cases where the model space is very large it can be burdensome to compute asymptotic quantities for each possible specification, put aside the more substantive and philosophical issue concerning the underpinnings of Bayesian inference.

Table 1: Prior Distributions

Parameters	*Priors*
β_{j,t_0}	$\sim N(\mu_j, \Sigma_j)$
σ^2	$\sim IG(\nu_{\Delta e}/2, s_{\Delta e}/2)$
q_j^2	$\sim IG(\nu_j/2, s_j/2)$
λ_j	$\sim Beta(a_j, b_j)$
π_j	$\sim Beta(c_j, d_j)$

3.3 Posterior Simulation

As already mentioned above, I follow Ravazzolo *et al.* (2007)'s procedure: a Gibbs sampler is used in combination with the data augmentation technique by Tanner and Wong (1987) for latent variables such as $S = (s_1, ..., s_m), B = \{\beta_t\}_{t=1}^T$ and $K = \{k_t\}_{t=1}^T$ where $\beta_t = (\beta_{0,t}, ..., \beta_{m,t})$ and $k_t = (k_{0,t}, ..., k_{m,t})$.

The complete data likelihood function is given by

$$\begin{aligned} p(\Delta e, B, K|x, \theta) \quad = \quad & \prod_{t=1}^{T} p(\Delta e_t|S, x_t, \beta_t, \sigma^2) \prod_{j=0}^{m} p(\beta_{jt}|\beta_{jt-1}, k_{jt}, q_j^2) \times \\ & \prod_{j=0}^{m} \pi_j^{k_{jt}} (1-\pi_j)^{1-k_{jt}} \end{aligned} \tag{6}$$

where $\Delta e = (\Delta e_1, ..., \Delta e_T)$ and $x = (x_1', ..., x_T')$ while the first two terms on the right hand side only entail evaluations of normal density functions.

Combining the prior and the data likelihood, one obtains the posterior density

$$p(\theta, S, B, K|\Delta e, x) \propto p(\theta)p(S)p(\Delta e, B, K|\theta, S, x) \tag{7}$$

where $\theta = (\pi_0, ..., \pi_m, q_0^2, ..., q_m^2, \sigma^2)$ collects the model parameters.

The Gibbs sampler consists of the following iterative steps

1. Draw S conditional on $B, K, \theta, \Delta e$ and x.

2. Draw K conditional on $S, \theta, \Delta e$ and x.

3. Draw B conditional on $S, K, \theta, \Delta e$ and x.

4. Draw θ conditional on $S, B, K, \Delta e$ and x.

The first step applies the the Kuo-Mallick (1998) algorithm, a simplified version of George and McCulloch (1993)'s. Each s_j is sampled from its conditional posterior distribution, $Pr(s_j|\Delta e, x, \theta, B, K, S_{-j})$, where $S_{-j} = (s_1, ..., s_{j-1}, s_{j+1}, ..., s_m)$. This distribution is Bernoulli $B(1, \widetilde{p}_j)$ with $\widetilde{p}_j = p_{j,1}/(p_{j,1}+ p_{j,0})$, where

$$p_{j,1} = \lambda_j \exp\{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(\Delta e_t - x_t B^{*}_{t,j})^2\}$$

and

$$p_{j,0} = (1-\lambda_j) \exp\{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(\Delta e_t - x_t B^{**}_{t,j})^2\}. \tag{8}$$

The vector $B^{*}_{t,j}$ corresponds to the vector $BS_t = (\beta_{1,t}s_1, ..., \beta_{m,t}s_m)'$ with its jth entry equal to $\beta_{j,t}$, while for $B^{**}_{j,t}$ the jth entry is 0.

The efficient sampling algorithm of Gerlach *et alia* (2000) is used in the second step, that is for structural breaks detection. The main advantage is in drawing k_{jt} without conditioning on the states β_{jt}, as Carter and Kohn (1994) instead do: the conditional posterior density for k_t, $t = 1, .., T$ unconditional on B is

$$\begin{aligned} p(k_t|K_{-t}, S, \theta, \Delta e, x) &\propto p(\Delta e|K, S, \theta, x)p(k_t|K_{-t}, S, \theta, x) \\ &\propto p(\Delta e^{t+1,T}|\Delta e^{1,t}, K, S, \theta, x) \\ &\quad\; p(\Delta e_t|\Delta e_{1,t-1}, k_{1,t-1}, S, \theta, x)p(k_t|K_{-t}, S, \theta, x). \end{aligned} \tag{9}$$

Gerlach *et alia* (2000) show how to evaluate the first two terms while the last one is obtained from the prior. When K_t and β_{jt} are higly dependent the sampler of Carter and Kohn (1994) breaks down completely: the higher the correlation (dependence), the bigger the efficiency gain using (9).

The third step applies the Kalman filter and smoother as in Carter and Kohn (1994) to derive the conditional mean and variance of the latent factors.

When a variable is not selected (i.e. $s_j = 0$), $k_{j,t}$ is drawn from the prior and $\beta_{j,t}$ unconditionally from the process in (5). Finally, the parameters in θ are easily sampled as independent conjugate priors are used and hence hyperparameters can be interpreted as pre-sample occurences.

The main goal is to produce measures of exchange rate dynamics forecast, Δe_{t+h} in (4) for $h \geq 1$, taking into account all features of uncertainty. The one-step ahead predictive density of Δe_{T+1} at time T given the sample observations is

$$p(\Delta e_{T+1}|\Delta e, x) = \int\int \sum_S \sum_K \sum_{K_{T+1}} p(\Delta e_{T+1}|S, \beta_{T+1}, \sigma^2)$$

$$p(\beta_{T+1}|\beta_T, K_{T+1}, q_0^2, ..., q_m^2) \prod_{j=0}^{m} \pi_j^{k_{j,T+1}} (1-\pi_j)^{1-k_{j,T+1}} p(B, K, S, \theta|\Delta e, x) dB d\theta, \tag{10}$$

where the averaging also considers the possibility of breaks at $T+1$ with weights given by $\prod_{j=0}^{m} \pi_j^{k_{j,T+1}} (1-\pi_j)^{1-k_{j,T+1}}$.[13] The draws from the sampling scheme at each step are used to simulate the distribution in (10).

4 Empirical Analysis

4.1 Data Description

Data are quarterly and cover the period from the first quarter of 1973 until the end of 2008: they have been collected from Datastream according to the sources specified in Wright.[14] Two major currencies are considered: the Canadian dollar and the Japanese yen, all *vis-a-vis* the US dollar. The dependent variable is the quarterly change in each bilateral nominal exchange rate, computed as end-of-period values.

As for the explanatory variables, I use a smaller set of variables ($m = 11$): a price level (CPI), a monetary aggregate (M1), real GDP, the annual inflation rate, the annual money growth rate,

[13]The formula above considers just the simple case of $h = 1$. Remember that in a direct forecasting setup, at each forecast horizon corresponds a different regression to be estimated.

[14]Use of quarterly data somehow contradicts the pure sense of "real-time" experiment as data are subject to revision and, hence, to Faust, Rogers and Wright (2003)'s criticism.

the annual growth rate in real income, a local stock market index (MSCI), the current account-GDP ratio, the short term (3-month) interest rate and the yield spread (10-year minus 3-month interest rates) and finally oil price.[15] Current account, GDP and money aggregates have been included upon seasonal adjustement unlike CPI inflation. For Japan, no appropriate data on the current account have been found for the sample under investigation, hence the relative exchange rate forecasts are based on ten potential predictors. All the $x's$ in (4) have been constructed as cross-country differentials by taking the log difference of all the aforementioned variables but interest rates and oil price. As mentioned in the previous section, they all have been demeaned in order for the intercept to vary over time regardless of movements in slope parameters.

4.2 In Sample Analysis

Prior to investigating the forecasting performance of the BMA approach, I shortly report a full sample (*ex post*) analysis. This is an interesting exercise in order to infer about the relevance of different model specifications and the evolution of coefficients over time.[16]

A first step is to evaluate the selection procedure in case of no structural breaks. Tables 2 and 3 report results with an almost noninformative prior for the error term and a moderately informative one for the coefficient variance.[17] For the Us dollar-Canadian dollar exchange rate, the number of models actually visited is big (about 763) denoting a high degree of uncertainty. Indeed the most likely model, including only the relative money growth rate, is assigned about two percent as posterior probability (first column in Table 2). This variable appears in most of the top ten models and it is the only one with a posterior inclusion probability greater than the prior probability of one half. The random walk model with drift receives adeguate support

[15]Wright (2008) considers a total number of fifteen variables but his simpler framework allows analytical derivation of posterior model probabilities and predictive means. In my case, a posterior simulator is needed and then in order to keep the analysis feasible a few variables have been discarded on the basis of prevailing practice in the literature. Finally, the spot oil price appeared in his working paper version and also in Rossi (2006) and has been kept here.

[16]In more details, I ran 15,000 iterations saving every other draw and burning the first one thousands. Convergence has been checked by both visual inspection of recursive plots and evaluation of Geweke's CD statistics for selected parameters.

[17]The latter choice is due to the fact that, in case of model selection for nested models, a less and less informative prior variance of the $\beta' s$, even keeping the same stand about the noise term, increasingly favors the restricted model, the random walk in this framework. This is likely related to the well-known *Lindley's paradox*.

(around 1.5 %) and belongs to the top part of the model ranking. Furthermore, the most visited models are evidently low-dimensional and they never display more than three regressors.

For the US dollar-Japanese yen exchange rate, the relative term spread emerges as a relevant predictor and outclasses the other explanatory variables. The most likely model which includes the relative term spread and the relative annual money growth has a probability of 2.5 percent. Again, very parsimonious specifications prevail but in this case the random walk with drift receives much less support.

Then dynamics in the parameters is introduced. Admittedly, the prior about both the variance in the transition equation and the break probabilities is not innocuous: yet, the influence is reflected mostly in the size of the coefficients while their time path is preserved. Inference about other quantities of interest (inclusion probabilities and top model ranking) appears not to be severely affected on average. Results in Tables 4 and 5 are based on the same values as in the case without breaks for the measurement equation variance and the initial condition of the slope coefficients. For the other parameters, it applies what said in section 4.2.

Overall, the amount of model uncertainty eases off quite a lot and the first ranked model turns out to be strongly favored by the sampler. Marginal inclusion probabilities change dramatically in some cases: the examples of the short rate differential (increase) and the spread differential (decrease) for the US dollar-Japanese yen exchange rate are telling. Similarly, for the US dollar-Canadian dollar it is the spot oil price and the relative current account over GDP that become relevant. In both cases the inclusion rates of price level related regressors have more than halved. The average effect is an increased parsimony in the most selected models with respect to the case of constant parameters.

The analysis in this section has confirmed that model uncertainty is undoubtedly important in the context of exchange rate modelling and is furthermore affected by the amount of flexibility each model is granted. When a fair number of traditional explanatory variables is considered, no single model stands out as the one with exclusive support from the data.

4.3 Out-of-Sample Forecasting Analysis

The aim of this paper is to ultimately test the forecasting performance of a flexible approach that combines plausible characteristics of economic variables. Several out-of-sample metrics are considered in order to assess the procedure proposed here. Recently, a surge of interest in the literature has concerned whether to use in sample or out of sample tests for predictability: the latter strategy is usually advocated to avoid data mining problems typical of the former approach. Inoue and Kilian (2004) question this view arguing that both of them are susceptible to data mining and should hence be corrected by using appropriate critical values. The BMA approach provides a further advantage in that it overcomes such problems deriving from reporting only "best results" by averaging over all possible combinations of explanatory variables and so data mining is not much of a concern for the present work.

A typical tool for forecast comparison (out-of-sample predictability) is the root-mean-square-prediction error (RMSPE) for the competing models, whose ratio is referred to as the Theil's U statistic: being the no-change model the null, a value less than one means that the model under scrutiny has more predictive power. However, it is often argued that such a measure may be misleading as it extremely penalizes single large errors while an appropriate decision rule should balance true costs and benefits. As a consequence, several metrics are usually considered altogether: among them, I look at the mean absolute error (MAE) and the Hit Rate. The latter is frequently used in financial applications as a weak proxy for a pure profit-based metric and indicates the fraction of times a given model correctly predicts the direction of change whereas the theoretical "hit sequence" for a random walk is one half.

In any case, a proper Bayesian analysis should care not only about the mean of the predictive distribution. In this respect, the Bayesian version of the Diebold Mariano statistic (henceforth, DMB), as illustrated in Koop *et al.* (2009), is applied. It essentially computes the average probability that the loss differential is unfavourable to the model under scrutiny: if this average probability is less than 0.5 then the BMA strategy has higher predictive power in this case. Moreover, Pesaran, Pettenuzzo and Timmermann (2006) point out that when the predictive densities are non-normal, the relative RMSPE is only slightly informative and thus suggest the predictive Bayes factor (henceforth, PBF) for pair-wise model comparison: it is a ratio between two different predictive likelihood (not *density*) functions or in other terms

$$BF_t = \frac{f_{RW}(r_t|...)}{f_{BMA}(r_t|...)} \tag{11}$$

where each $f_*(r_t|...)$ stands for the empirical probability density function derived from the simulated predictive density by using a kernel estimator, the Epanechinov kernel in this case. A value greater than one supports the predictive performance of the benchmark model: this calculation is carried out for each observation in the recursive forecasting procedure and the average value is finally considered.[18]

The forecast comparison starts from 1990:Q1 and includes two subsets: the main one, ending in 2008:Q4, which consists of 76 out-of-sample observations and a shorter one (up to 2005:Q4) in conformity with Wright. Results are shown in Tables 6 and 8 while Tables 7 and 9 accurately list the prior settings of each forecasting model.

For the US dollar-Canadian dollar case, the verdict for point forecasts is mixed. The most used indicator, the RMSE, always favor the no-change model: in the very few cases in which a "BMA plus break" specification delivers a smaller bias, the standard deviation of the resulting forecast errors increases to the extent that it more than offset the former gain. Mixture innovation models on average perform better than models with smooth time variation. The mean absolute error (MAE) instead tells a slightly different story as model averaging, both alone (always) and together with structural instability (in a couple of cases, namely when the variance of the initial condition of the slopes is strongly reduced), is capable to outperform the random walk model. Interestingly, fixed-parameter models all have a negative bias that is larger in absolute value than the benchmark's: it implies that their (point) forecast error distribution is higly asymmetric with many small positive and large negative values.[19] The Hit Rate, which is regarded as a better economic criterion, is always greater than one half apart from the majority

[18] Alternatively, Geweke and Amisano (2009) consider the sum of log Bayes factors. In this way, one is also able to single out the observations responsible for the final outcome.

[19] The comparison of relative standard deviations from the left column in Table 6 further confirms this intuition. The ultimate proof should of course come from the plot of such error distributions.

of TVP specifications: in this respect, it seems easier to do better than just tossing a coin.

It is worth noting that, among the specifications allowing for time variation tested in the paper, TVP models fare decidedly worse by any metric. Instead shrinkage, that ultimately implies controlling the amount of prior variability for the initial conditions of the β's, again proves quite an effective tool (look for instance at BMA-B4 *vs.* BMA-B2).[20] Contrary to Wright, favoring more parsimonious models by imposing λ equal to 0.2 does not necessarily pay off in this case.

Finally, in the shorter subsample (1990:Q1 - 2005:Q4) a few specifications do beat the random walk according to the RMSE statistic: these are namely the models allowing for (but not imposing) time variaton at each time.[21] By and large, the performance of all specifications in Table 6 strictly improves in the sample analysed in Wright. In more general terms, Figure 2 shows that the size of forecast errors has greatly increased since 2003, irrespective of the underlying model. This leads to think of some new factor(s) that has(have) been driving the dynamics of the USD-CAD exchange rate ever since.[22]

For the US Dollar-Japanese Yen, the evidence is slightly in favor of the BMA approach. As for point predictions, averaging over fixed-parameter models outclasses the no-change model in terms of both RMSE and MAE: interestingly, the average bias is negative as well as in the previous case. In general, the forecast errors generated by the benchmark model are widely dispersed and almost all the competitors tried do better in this respect even though the introduction of structural instability erodes the overperformance, particularly in the presence of continuous time-variation. Within each subclass (constant parameter, mixture innovation and TVP) of models there seems to exist a positive relationship between bias and standard deviation of forecast errors.

Shrinkage in terms of model dimension, controlled by the λ parameter, seems to work unlike the US Dollar-Canadian dollar case. A plausible conjecture is that this prior helps in focusing

[20] Strictly speaking, the level of shrinkage here is not directly comparable to Wright that uses a different prior for the slope coefficients. For the sake of precision, the way time variation is modelled in the present work makes each β a state variable for which only the initial condition has to be specified.

[21] They are labelled as BMA-B in Table 6. Results are not reported to save space but are available upon request.

[22] It might also have something to do with the status of the Canadian dollar as a "commodity currency". Anyway, this interesting aspect is left unexplored in this draft.

on the very best models that happen to be exactly the most parsimonious ones. Furthermore, the proposed mixing strategy consistently deliver better signals of future change direction in the value of the currency, regardless of how parameters are modelled.

Restricting the out-of-sample time span to the 2005:Q4 does not affect the response even though the overall BMA outperformance deteriorates. The relevant difference is just that the only specification with breaks (BMA-B6) capable to beat the random walk in the longer sample is now outperformed. Its squared forecast errors are plotted in Figure 4 together with the benchmark's and the best constant-parameter BMA's. In general, the onset of the financial cirsis exerts a dramatic effect on the size of forecast errors of every model. Unlike the USD-CAD exchange rate, there are a couple of episodes, in the middle and late nineties, of comparable or even greater impact.

In his final conclusion Wright recognizes two crucial aspects to keep in mind:

> (Results) are all very much consistent with the idea that model and parameter uncertainty are the stumbling blocks to exchange rate forecasting (given that the exchange rate is so close to being a random walk), and that the researcher who wants to get good out-of-sample prediction, rather than in-sample fit, should use shrinkage methods.[23]

As a whole, I find not dissimilar results when exactly the same sample as his is considered. Yet, in the long out-of-sample exercise a random walk model for the US dollar-Canadian dollar becomes difficult to improve upon.

The introduction of structural instability hardly appears beneficial. Jochmann, Koop and Strachan (2008) come to the same conclusion with a similar approach to predict macro variables. However, in the present study the mixture innovation approach *does* bring some improvements for one currency provided that substantial prior mass is put on the occurence of few(er) rather than continuous breaks. This finding is extremely compelling if coupled with the influential work of Stock and Watson (1996) who report that constant parameter models often beat TVP

[23] Likewise Pettenuzzo and Timmermann (2005) claim that shrinkage results in better forecasting performance with parameter and model uncertainty and also in presence of structural breaks, when these breaks are few and their distribution reasonably precisely estimated.

models in terms of predictions in spite of convincing evidence of instability found in their wide analysis of US macro variables. Extending finally the focus from point to distribution forecasts, models subject to breaks have the best scores for both currencies: in particular, for the US dollar-Canadian dollar they are the only specifications able to beat the benchmark. All in all, the mixture innovation approach seems to be a fruitful avenue to pursue in empirical applications dealing with instability.

5 Conclusions

The last years have witnessed a renewed interest in the topic of exchange rate forecasting. New models, new methods and new inference tools have partially disrupted the daunting primacy of a no change model. For instance, Wright (2008) suggests the potential of Bayesian model averaging as forecasts are combined in a sensible way. The present paper checks in a slightly different framework whether the latter finding carries over to a longer sample and more importantly tests the marginal contribution of modelling structural instability as well. The focus is on one quarter ahead predictions, a short term horizon that is well known to be the most severe hurdle to pass for fundamental-based models.

The results in the previous sections show that Bayesian Model Averaging is an useful technique for predicting the movements in the US dollar-Japanese yen exchange rate. Evidence for the US dollar-Canadian dollar case is instead mixed as BMA is superior to the benchmark only in the subsample, originally studied by Wright, ending in 2005. The final verdict depends on which criterion informs the evaluation process, the more favourable being the Hit rate and, to a lesser extent, the MAE.

The novelty with respect to Wright (2008) is the introduction of parameter instability in quite a flexible way, by means of a mixture innovation model. While this approach rewards in sample by notably reducing the amount of model uncertainty, in the out-of-sample comparison it hardly improves upon a model averaging approach with just constant parameters. The higher flexibility offered by this approach may be dominated by the additional amount of estimation error/uncertainty in the parameters. Promising results are, anyway, obtained when “favouring” rare breaks through prior information. In this sense, the mixture innovation model appears superior to the more commonly used TVP model, which is restrictive in that variables are forced to undergo continuous shifts in the parameters.

Giordani and Villani (2009) offer a wider investigation of models that incorporate instabilities for macro variables in a very similar way as here. In particular, they allow for further levels of complexity on the grounds that outliers or decreases in variance may be mistaken for changes in the conditional mean, particularly in application with high variability. Tight priors in my case may be limiting this shortcoming. In this respect, I plan to also add stochastic volatility and additive outliers.[24]

A further aspect worth exploring is the effect of the random walk structure imposed in the case of variation on the predictive performance. D'Agostino, Gambetti and Giannone (2009) point out the role that a stability constraint plays in ruling out excessively volatile behaviour in multi-horizon forecasts, a concern that is anyway less warranted here with one-quarter ahead forecasts.[25] Finally, it could be explored a larger information set including other variables that have recently been shown to have predictive content, such as measures of external imbalance.

Anyway, both Wright (2008) and Engel, West and Mark (2007) point out that we should not expect exchange rates to behave quite differently from a random walk, so gains from any procedure alternative to a simple "no-change forecast" strategy are likely to be moderate.

References

[1] Altavilla C. and De Grauwe P., "Forecasting and Combining Competing Models of Exchange Rate Determination", *Applied Economics*, 2010, forthcoming.

[24] At higher frequencies the literature routinely employs models with variation in the conditional second moment. Among the recent studies, Della Corte, Sarno and Tsiakas (2009) are very close in spirit to the present paper as they test the profitability of portfolio allocation strategies based on Bayesian model averaging. The models exploit the deviation from equilibrium levels implied by typical macro fundamentals and forward premia: notably, specifications featuring stochastic volatility prevail regardless of the specific strategy or fundamental being used.

[25] On the other hand, the same point is more relevant if one think of the trajectory of the estimated coefficients in the "pseudo-in sample" analysis before each prediction. In this sense, my best results are easily explained as the strong priors on break frequency and size serve the purpose of making the system less unstable. Then, which structure has to be preferred remains debatable.

[2] Avramov D., "Stock Return Predictability and Model Uncertainty", *Journal of Financial Economics*, 2002, N.64, 423-458.

[3] Bacchetta P., van Wincoop E. and Beutler T., "Can Parameter Instability Explain the Meese-Rogoff Puzzle?", *Mimeo*, 2009.

[4] Bates J.M. and Granger C., "The Combination of Forecasts", *Operations Research Quarterly*, 1969, N.20, 451-468.

[5] Brock W., Durlauf S. and West K., "Model Uncertainty and Policy Evaluation: Some Theory and Empirics", *Journal of Econometrics*, 2006, N.136, 629-664.

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

[7] Clark T. E. and McCracken M., "The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence", *Journal of Money, Credit, and Banking*, 2006, V. 38, No. 5, 1127-1148.

[8] Clark T. E. and West K. D., "Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis", *Journal of Econometrics*, 2006, V.127, 155-186.

[9] Cremers K.J., "Stock Return Predictability: A Bayesian Model Selection Perspective", *Review of Financial Studies*, 2002, N.15(4), 1223-1249.

[10] Cuaresma J. and Doppelhofer G., "Nonlinearities in Cross-Country Growth Regressions: A Bayesian Averaging of Thresholds (BAT) Approach", *University of Vienna Working Paper*, 2006, N.0608.

[11] D'Agostino A., Gambetti L. and Giannone D., "Macroeconomic Forecasting and Structural Change", *ECARES Working Paper*, 2009, N.2009-020.

[12] Della Corte P., Sarno L. and Tsiakas I., "An Economic Evaluation of Empirical Exchange Rate Models: Robust Evidence of Predictability and Volatility Timing", *Review of Financial Studies*, 2009, 22, 3491-3530.

[13] Doppelhofer G., Miller R.I. and Sala-i-Martin X., "Determinants of Long Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach", *NBER Working Paper*, 2000, N.7750.

[14] Draper D., "Assessment and Propagation of Model Uncertainty", *Journal of the Royal Statistical Society*, 1995, 4597.

[15] Engel C., Mark N. and West K., "Exchange Rate Models Are Not as Bad as You Think", *NBER Macroeconomics Annual*, 2007, 381-441.

[16] Eklund J. and Karlsson S., "Forecast Combination and Model Averaging Using Predictive Measures", *Econometric Reviews*, 2007, V.26, 2-4, 320 -363.

[17] Faust J., Rogers J.H. and Wright J.H., "Exchange Rate Forecasting: The Errors We've Really Made", *Journal of International Economics*, 2003, N.60, 35-59.

[18] Fernandez C., Ley E. and Steel M.F.J., "Model Unceratinty in Cross-Country Growth Regressions", *Journal of Applied Econometrics*, 2001, N.16, 563-576.

[19] George, E.I. and McCulloch R. E., "Variable selection via gibbs sampling", *Journal of the American Statistical Association*, 1993, N.88, 881-889.

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

[21] Geweke J. and Amisano G., "Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns", *International Journal of Forecasting*, forthcoming.

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

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

[24] Giordani P. and Villani M., "Forecasting Macroeconomic Time Series With Locally Adaptive Signal Extraction", *Sveriges Riksbank Working Paper Series*, 2009, N.234.

[25] Gourinchas P. O. and Rey H., "International financial adjustment". *NBER Working Paper*, 2005, No. 11155.

[26] Guidolin M. and Na C.F., "The Economic and Statistical Value of Forecast Combinations under Regime Switching: An Application to Predictable US Returns", Rapach D. and Wohar M. (eds.), "Frontiers of Economics and Globalization" Volume 3, *Forecasting in the Presence of Structural Breaks and Uncertainty*, Emerald Publishing Ltd. & Elsevier Press, 2007, pp. 595-657.

[27] Inoue A. and Kilian L., "In-sample or Out-of-sample Tests of Predictability: Which One Should We Use?", *Econometrics Review*, 2004, N.23, 371-402.

[28] Inoue A. and Kilian L., "How Useful is Bagging in Forecasting Economic Time Series? A Case Study of US CPI Inflation", *Journal of the American Statistical Association*, 2008, N.103, 511-522.

[29] Jacobson T. and Karlsson S., "Finding Good Predictors for Inflation. A Bayesian Model Averaging Approach", *Journal of Forecasting*, 2004, N.23, 479-496.

[30] Jochmann M., Koop G. and Strachan R.W., "Bayesian Forecasting using Stochastic Search Variable Selection in a VAR Subject to Breaks", *International Journal of Forecasting*, 2010, forthcoming.

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

[32] Koop G. and Potter S., "Are Apparent Findings of Nonlinearity Due to Structural Change?", *The Econometrics Journal*, 2001, N.4, 37-55.

[33] Koop G. and Potter S., "Forecasting in Dynamic Factor Models Using Bayesian Model Averaging", *The Econometrics Journal*, 2006.

[34] Koop G., Leon-Gonzalez R. and Strachan R., "On the transmission of the monetary policy mechanism", *Journal of Economic Dynamics and Control*, 2009, 33, 997-1017.

[35] Koop G., Garratt T., Mise E. and Vahey S., "Real-time Prediction with UK Monetary Aggregates in the Presence of Model Uncertainty", *Journal of Business and Economic Statistics*, 2009, 27, 480-491.

[36] Kuo L. and Mallick, B., "Variable Selection for Regression Models", *Sankhya*, 1998, 60, 65-81.

[37] Leamer E. E., *Specification Searches*, New York: Wiley, 1978.

[38] Mark N.C., "Exchange Rates and Fundamentals. Evidence on Long Horizon Predictability", *American Economic Review*, 1995, N.85, 201-218.

[39] Meese R.A. and Rogoff K., "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?", *Journal of International Economics*, 1983, N.14, 3-24.

[40] Molodtsova T. and Papell D.H., "Out-of-sample Exchange Rate Predictability with Taylor Rule Fundamentals", *Journal of International Economics*, 2009, 77, 167-180.

[41] Neely C. and Sarno L., "How Well Do Monetary Fundamentals Forecast Exchange Rates?", *Federal Reserve Bank of St. Louis Review* , 2002, issue September, 51-74.

[42] Neely C. and Weller P., "Predictability in International Asset Returns: A Reexamination", *The Journal of Financial and Quantitative Analysis*, 2000, V.35, N.4, 601-620.

[43] Pesaran H., Pettenuzzo D. and Timmermann A., "Forecasting Time Series Subject to Multiple Structural Breaks", *Review of Economic Studies*, 2006, V. 73, 10057-1084.

[44] Pettenuzzo D. and Timmermann A., "Predictability of Stock Returns and Asset Allocation under Structural Breaks", *University of California San Diego Working Paper*, 2005.

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

[46] Raftery A., Madigan D. and Hoeting J., "Bayesian Model Averaging for Linear Regression Models", *Journal of the American Statistical Association*, 1997, Vol. 92, N.437, 179-191.

[47] Raftery A., Madigan D. and Hoeting J., "A Tutorial for Bayesian Model Averaging", 1999.

[48] Rapach D. and Wohar M., "In-sample vs Out-of-sample Tests of Stock Return Predictability in The Context of Data Mining", *Journal of Empirical Finance*, 2006, N.13, 231-247.

[49] 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.), "Frontiers of Economics and Globalization" Volume 3, *Forecasting in the Presence of Structural Breaks and Uncertainty*, Emerald Publishing Ltd. & Elsevier Press, 2007, pp. 561-594.

[50] Rossi B., "Are Exchange Rates Really Random Walks? Some Evidence Robust to Parameter Instability", *Macroeconomic Dynamics*, 2006, 10, 20-38.

[51] Rogoff K. and Stavrakeva V., "The Continuing Puzzle of Short Horizon Exchange Rate Forecasting", *NBER Working Paper*, 2008, N. W14071.

[52] Sarno L. and Valente G., "Exchange Rates ad Fundamentals: Footloose or Evolving Relationship?", *Journal of the European Economic Association*, forthcoming.

[53] Schinasi G. and Swamy P., "The Out-of-Sample Forecasting Performance of Exchange Rate Models When Coefficients Are Allowed to Change", *Journal of International Money and Finance*, 1989, 8, 375-390.

[54] Stock J. and Watson W., "Evidence on Structural Instability in Macroeconomic Time Series Relations", *Journal of Business and Economic Statistics*, 1996, N. 14, 11-30.

[55] Stock J. and Watson W., "An Empirical Comparison of Methods for Forecasting Using Many Predictors", *Manuscript*, 2005.

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

[57] Timmermann A., "Forecast Combinations", in G. Elliott, CWJ Granger and A. Timmermann, eds., *Handbook of Economic Forecasting*, 2005, Amsterdam: North Holland.

[58] Wright J. H., "Bayesian Model Averaging and Exchange Rate Forecasts", *Journal of Econometrics*, 2008, N. 146, 329-341.

[59] Wright J. H., "Forecasting U.S. Inflation by Bayesian Model Averaging", *Journal of Forecasting*, 2009, N. 28, 131-144.

Table 2: US dollar- Canadian dollar most likely models - No Breaks

P. Model Prob.	*y-y**	*m-m**	*p-p**	*msci-msci**	*(ca/y)-(ca*/y*)*	*i-i**	*spread*	$\Delta_{12}y - \Delta_{12}y^*$	$\Delta_{12}m - \Delta_{12}m^*$	$\Delta_{12}p - \Delta_{12}p^*$	*oil*
$p = 0,0203$	0	0	0	0	0	0	0	0	1	0	0
$p = 0,0194$	0	0	1	0	0	0	0	0	1	1	0
$p = 0,0191$	0	0	1	0	0	0	0	0	1	0	0
$p = 0,0191$	0	0	0	0	0	0	0	0	1	1	0
$p = 0,0158$	0	0	1	0	0	0	0	0	0	1	0
$p = 0,0155$	0	0	0	0	0	0	0	0	0	1	0
$p = 0,0154$	0	0	0	0	0	0	0	0	0	0	0
$p = 0,0152$	0	0	1	0	0	0	0	0	0	0	0
$p = 0,0145$	0	1	1	0	0	0	0	0	1	0	0
$p = 0,0138$	0	1	0	0	0	0	0	0	1	0	0
P. Incl. Prob.	$0,286$	$0,399$	$0,496$	$0,391$	$0,059$	$0,124$	$0,159$	$0,128$	$0,522$	$0,493$	$0,09$

Notes: The table lists the ten models with the highest posterior probabilities as selected by the Gibbs sampler. The sequence of 0 and 1 denotes the resulting model specifications: the intercept is always included and so not reported here.
Finally, the first column shows the posterior probability associated to each selected model while the last row the posterior inclusion probability of each potential predictor.

Table 3: US dollar- Japanese yen most likely models No breaks

P. Model Prob.	*y-y**	*m-m**	*p-p**	*msci-msci**	*i-i**	*spread*	$\Delta_{12}y - \Delta_{12}y^*$	$\Delta_{12}m - \Delta_{12}m^*$	$\Delta_{12}p - \Delta_{12}p^*$	*oil*
$p = 0,025$	0	0	0	0	0	1	0	1	0	0
$p = 0,0235$	0	0	0	0	0	1	0	0	0	0
$p = 0,022$	0	1	0	0	0	1	0	0	0	0
$p = 0,021$	0	0	1	0	0	1	0	0	0	0
$p = 0,0198$	0	0	0	1	0	1	0	0	0	0
$p = 0,016$	0	1	1	0	0	1	0	0	0	0
$p = 0,0155$	0	0	1	0	0	1	0	1	0	0
$p = 0,0152$	0	1	0	0	0	1	0	1	0	0
$p = 0,0132$	0	1	0	1	0	1	0	0	0	0
$p = 0,0128$	0	0	1	0	0	0	0	0	0	0
P. Incl. Prob.	$0,2119$	$0,425$	$0,441$	$0,378$	$0,211$	$0,598$	$0,1592$	$0,414$	$0,136$	$0,203$

Notes: The table lists the ten models with the highest posterior probabilities as selected by the Gibbs sampler. The sequence of 0 and 1 denotes the resulting model specifications: the intercept is always included and so not reported here.
Finally, the first column shows the posterior probability associated to each selected model while the last row the posterior inclusion probability of each potential predictor.

Table 4: US dollar- Canadian dollar most likely models - With Breaks

P. Model Prob.	*y-y**	*m-m**	*p-p**	*msci-msci**	*(ca/y)-(ca*/y*)*	*i-i**	*spread*	$\Delta_{12}y - \Delta_{12}y^*$	$\Delta_{12}m - \Delta_{12}m^*$	$\Delta_{12}p - \Delta_{12}p^*$	*oil*
$p = 0,129$	0	0	0	0	0	0	0	0	0	0	0
$p = 0,05$	0	0	1	0	0	0	0	0	1	0	0
$p = 0,042$	0	1	0	0	0	0	0	0	0	0	0
$p = 0,032$	1	0	0	0	0	0	0	0	0	0	0
$p = 0,031$	0	0	0	0	0	0	0	0	0	1	0
$p = 0,03$	0	0	1	0	0	0	0	0	0	0	0
$p = 0,028$	0	0	0	0	1	0	0	0	0	0	0
$p = 0,028$	0	0	0	1	0	0	0	0	0	0	0
$p = 0,02$	0	0	0	0	0	0	0	0	0	0	1
$p = 0,014$	0	0	0	0	1	0	0	0	0	0	1
P. Incl. Prob.	0,212	0,226	0,197	0,161	0,216	0,08	0,065	0,075	0,256	0,196	0,219

Notes: The table lists the ten models with the highest posterior probabilities as selected by the Gibbs sampler. The sequence of 0 and 1 denotes the resulting model specifications: the intercept is always included and so not reported here.
Finally, the first column shows the posterior probability associated to each selected model while the last row the posterior inclusion probability of each potential predictor.

Table 5: US dollar- Japanese yen most likely models with Breaks

P. Model Prob.	*y-y**	*m-m**	*p-p**	*msci-msci**	*i-i**	*spread*	$\Delta_{12}y - \Delta_{12}y^*$	$\Delta_{12}m - \Delta_{12}m^*$	$\Delta_{12}p - \Delta_{12}p^*$	*oil*
$p = 0,13$	0	0	0	0	1	0	0	1	0	0
$p = 0,0736$	0	0	0	0	1	0	0	0	0	0
$p = 0,0635$	0	0	0	1	1	0	0	1	0	0
$p = 0,0358$	1	0	0	0	1	0	0	1	0	0
$p = 0,0343$	0	0	0	1	1	0	0	0	0	0
$p = 0,0314$	0	0	1	0	1	0	0	1	0	0
$p = 0,0309$	0	1	0	0	1	0	0	1	0	0
$p = 0,02954$	0	0	0	0	1	0	0	1	0	1
$p = 0,02077$	1	0	0	0	1	0	0	0	0	0
$p = 0,01923$	0	0	0	0	1	1	0	1	0	0
P. Incl. Prob.	$0,196$	$0,1798$	$0,192$	$0,306$	$0,893$	$0,181$	$0,071$	$0,628$	$0,0537$	$0,155$

Notes: The table lists just the ten models with the highest posterior probabilities as selected by the Gibbs sampler. The sequence of 0 and 1 denotes the resulting model specifications: the intercept is always included and so not reported here.
The first column shows, finally, the posterior probability associated to each selected model.

Table 6: US dollar- Can. dollar – Forecast Evaluation

	Theil's U	*MAE*	*Hit Rate*	*Predictive Bayes Factor*	*Bayesian DM*	*Bias*	*St. Dev.*
RW	1	2.23	0.5	1		0.0796	3.2462
BMA-NB1	1.0128	2.226	0.592	1.015	0.506	−0.108	3.287
BMA-NB2	1.008	2.212	0.579	1.012	0.5046	−0.145	3.27
BMA-NB3	1.009	2.216	0.5658	1.0108	0.5032	−0.1532	3.272
BMA-NB4	1.008	2.212	0.5658	1.0048	0.5036	−0.147	3.2699
BMA-NB5	1.0051	2.208	0.5526	1.0183	0.5032	−0.1503	3.2602
BMA-NB6	1.0028	2.207	0.5526	1.0188	0.5025	−0.119	3.2543
BMA-B1	1.0673	2.44	0.5395	1.1498	0.545	0.629	3.4082
BMA-B2	1.0036	2.2786	0.5263	1.04	0.5068	0.396	3.2349
BMA-B3	1.0086	2.29426	0.5395	1.0435	0.509	0.4737	3.2407
BMA-B4	1.0021	2.249	0.5263	1.02	0.4977	0.1967	3.2481
BMA-B5	1.0547	2.4207	0.5658	0.9895	0.5391	0.4898	3.3896
BMA-B6	1.0047	2.2216	0.5789	1.1341	0.5003	0.0286	3.266
BMA-B7	1.0073	2.2262	0.5658	0.9917	0.5004	0.009	3.2708
BMA-B8	1.0118	2.2546	0.5789	0.9909	0.5017	0.107	3.2839
BMA-BTVP1	1.2239	2.9579	0.4605	1.409	0.6077	1.8421	3.5215
BMA-BTVP2	1.1094	2.6118	0.4737	1.118	0.5546	1.1982	3.3973
BMA-BTVP3	1.0445	2.43	0.4342	1.1184	0.5362	0.8549	3.282
BMA-BTVP4	1.0551	2.4076	0.5658	1.0076	0.5248	0.5952	3.3741

Notes: The table reports the forecasting performance of the Bayesian Model Averaging approach in both absence (NB) and presence (B) of structural breaks with several prior hyperparameters as detailed in Table 7. The Predictive Bayes Factor is computed as a ratio of predictive likelihoods, namely $\frac{f_{RW}(r_t|...)}{f_{BMA}(r_t|...)}$, where the no-change model, denoted as RW, is the benchmark. Bias and Standard Deviation in the last two columns are such that $RMSFE \equiv [Bias]^2 + [SD]^2$.

Table 7: US dollar - Can. dollar – Prior hyperparameters for forecast evaluation

	Inclusion prob.	*Slope coeffs*	*Residual variance*	*Break prob.*	*State eq. variance*
BMA-NB1	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 5$	$\nu1 = 2$ $s1 = \nu1 * 7$		
BMA-NB2	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 1$	$\nu1 = 2$ $s1 = \nu1 * 7$		
BMA-NB3	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$		
BMA-NB4	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 1$	$\nu1 = 2$ $s1 = \nu1 * 7$		
BMA-NB5	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$		
BMA-NB6	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.1$	$\nu1 = 2$ $s1 = \nu1 * 7$		
BMA-B1	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 5$ $d = 95$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B2	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 5$ $d = 95$	$\nu2 = 30$ $s2 = \nu2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B3	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 1$ $d = 19$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B4	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.1$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 5$ $d = 95$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B5	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 1$ $d = 19$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B6	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.1$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 5$ $d = 95$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B7	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 5$ $d = 95$	$\nu2 = 30$ $s2 = \nu2 * 0.001^2 * (1/0.05) * V(\beta_0)$
BMA-B8	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 5$ $d = 95$	$\nu2 = 30$ $s2 = \nu2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-TVP1	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 1$ $d = 0$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * V(\beta_0)$
BMA-TVP2	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 1$ $d = 0$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * V(\beta_0)$
BMA-TVP3	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 1$ $d = 0$	$\nu2 = 30$ $s2 = \nu2 * 0.01^2 * V(\beta_0)$
BMA-TVP4	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.1$	$\nu1 = 2$ $s1 = \nu1 * 7$	$c = 1$ $d = 0$	$\nu2 = 2$ $s2 = \nu2 * 0.01^2 * V(\beta_0)$

Table 8: US dollar - Jap. yen – Forecast Evaluation

	Theil's U	*MAE*	*Hit Rate*	*Predictive Bayes Factor*	*Bayesian DM*	*Bias*	*St. Dev.*
RW	1	4.2607	0.5	1		−0.597	5.5718
BMA-NB1	0.9785	4.1842	0.5789	0.9838	0.4989	0.0125	5.483
BMA-NB2	0.9771	4.16	0.579	0.9568	0.4958	−0.092	5.475
BMA-NB3	0.9751	4.1478	0.566	0.964	0.495	−0.1815	5.461
BMA-NB4	0.984	4.219	0.5395	0.983	0.498	−0.097	5.513
BMA-NB5	0.9813	4.193	0.566	0.987	0.4982	−0.239	5.494
BMA-B1	1.0997	4.8905	0.5263	0.9912	0.5187	2.0964	5.7949
BMA-B2	1.0809	4.8078	0.5789	0.9344	0.5065	2.28	5.6104
BMA-B3	1.079	4.7983	0.5789	0.9366	0.5062	2.286	5.59
BMA-B4	1.0184	4.4123	0.57895	0.9822	0.4958	1.5532	5.491
BMA-B5	1.019	4.4114	0.57895	0.9173	0.4957	1.5456	5.497
BMA-B6	0.993	4.32	0.5658	0.9591	0.4966	1.051	5.46
BMA-B7	1.0496	4.63	0.5263	0.9509	0.5068	1.98	5.54
BMA-B8	1.0495	4.616	0.5526	0.9472	0.5063	1.9815	5.5375
BMA-TVP1	1.2223	5.55	0.5395	1.065	0.5534	2.9525	6.1868
BMA-TVP2	1.1569	5.11	0.5263	1.026	0.531	2.4541	6.001

Notes: The table reports the forecasting performance of the Bayesian Model Averaging approach in both absence (NB) and presence (B) of structural breaks with several prior hyperparameters as detailed in Table 9. The Predictive Bayes Factor is computed as a ratio of predictive likelihoods, namely $\frac{f_{RW}(r_t|\ldots)}{f_{BMA}(r_t|\ldots)}$, where the no-change model, denoted as RW, is the benchmark. Bias and Standard Deviation in the last two columns are such that $RMSFE \equiv [Bias]^2 + [SD]^2$.

Table 9: US dollar - Jap. yen – Prior hyperparameters for forecast evaluation

	Inclusion prob.	*Slope coeffs*	*Residual variance*	*Break prob.*	*State eq. variance*
BMA-NB1	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$		
BMA-NB2	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 1$	$\nu 1 = 2$ $s1 = \nu 1 * 30$		
BMA-NB3	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$		
BMA-NB4	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 1$	$\nu 1 = 2$ $s1 = \nu 1 * 30$		
BMA-NB5	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$		
BMA-B1	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 2$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B2	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 2$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B3	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 30$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B4	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.1$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 2$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B5	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.1$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 30$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B6	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.1$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 2$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B7	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 2$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-B8	$\lambda = 0.2$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 5$ $d = 95$	$\nu 2 = 30$ $s2 = \nu 2 * 0.01^2 * (1/0.05) * V(\beta_0)$
BMA-TVP1	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 1$ $d = 0$	$\nu 2 = 30$ $s2 = \nu 2 * 0.01^2 * V(\beta_0)$
BMA-TVP2	$\lambda = 0.5$	$\beta_0 = 0$ $V(\beta_0) = 0.5$	$\nu 1 = 2$ $s1 = \nu 1 * 30$	$c = 1$ $d = 0$	$\nu 2 = 30$ $s2 = \nu 2 * 0.01^2 * V(\beta_0)$

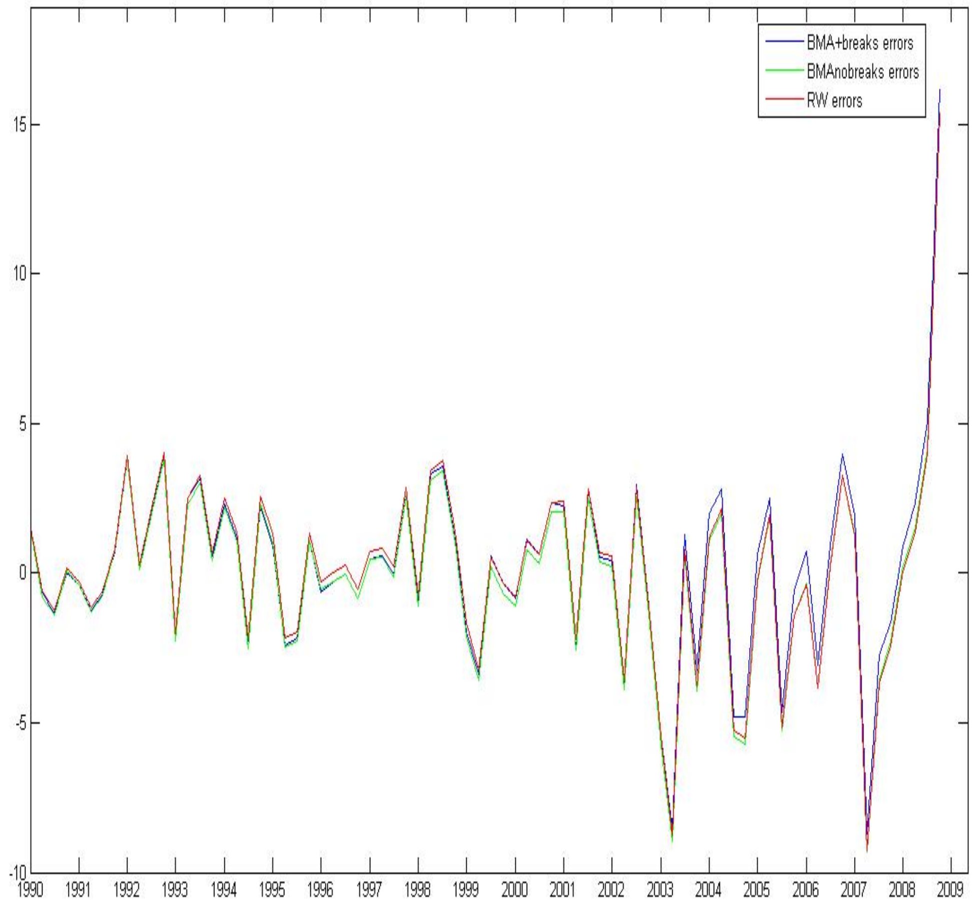

Figure 1: US Dollar vs Canadian Dollar - Forecast errors

Notes: The graph above plot the forecast errors of the no-change model and the best specifications using the BMA approach with constant parameters and structural instability, respectively.

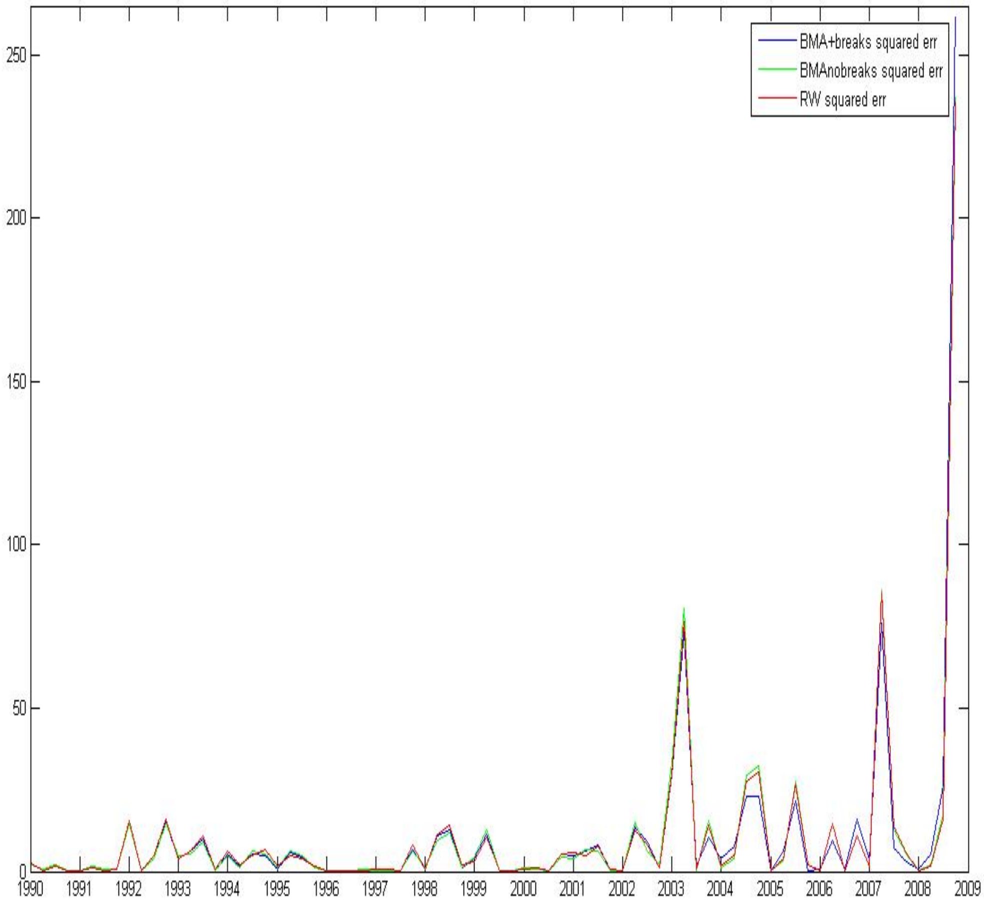

Figure 2: US Dollar vs Canadian Dollar - Squared forecast errors

Notes: The graph above plot the forecast errors of the no-change model and the best specifications using the BMA approach with constant parameters and structural instability, respectively.

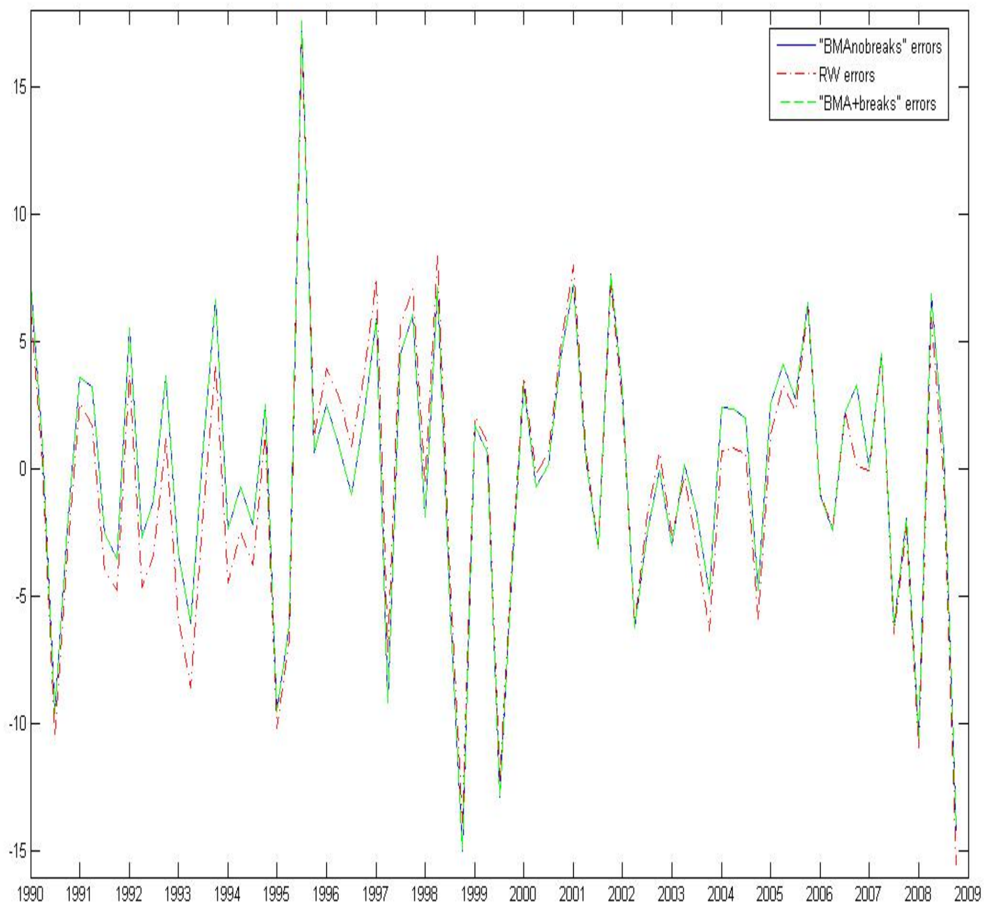

Figure 3: US Dollar vs Japanese Yen - Forecast errors

Notes: The graph above plot the forecast errors of the no-change model and the best specifications using the BMA approach with constant parameters and structural instability, respectively.

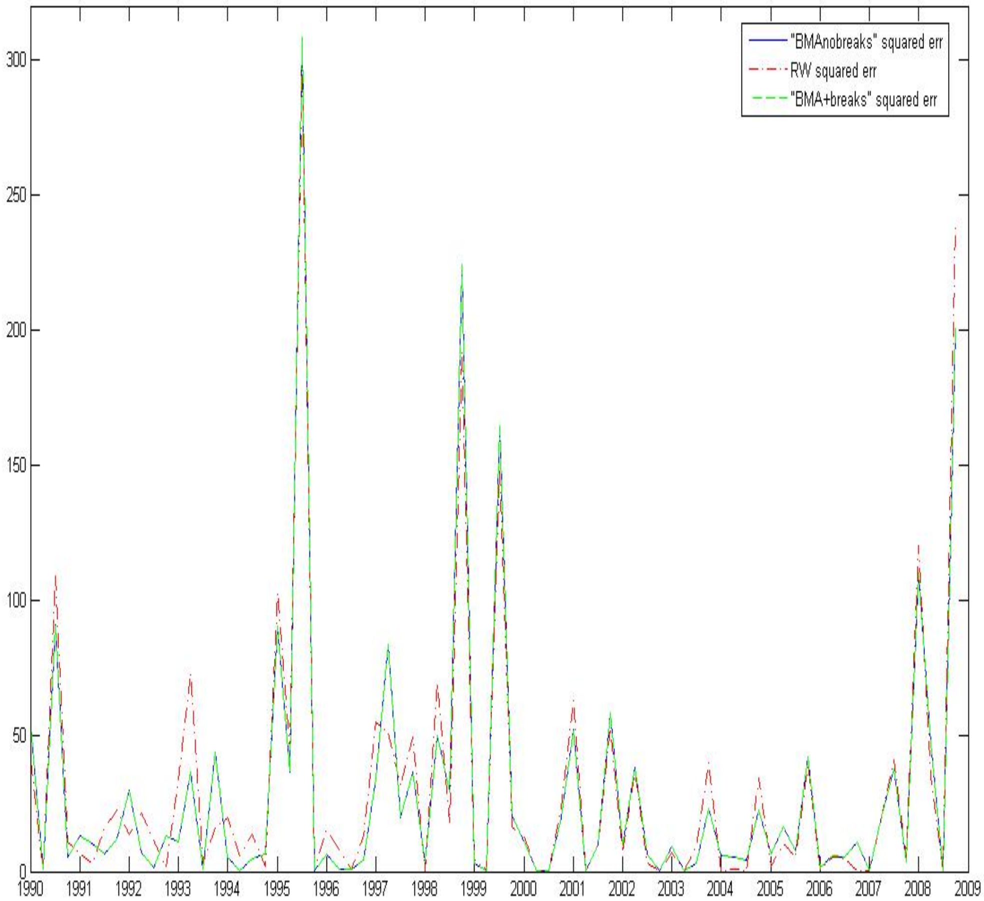

Figure 4: US Dollar vs Japanese Yen - Squared forecast errors

Notes: The graph above plot the forecast errors of the no-change model and the best specifications using the BMA approach with constant parameters and structural instability, respectively.

Bayesian Modelling of Time-Varying Prices and Quantities of Risk: Evidence from the US REIT Market*

Abstract

In this paper we analyze and compare the asset pricing properties and dynamics of publicly traded real estate to general asset classes like stocks and bonds. We extend the multifactor pricing model in Karolyi and Sanders (1998) with time-varying factor sensitivities and risk premia to a Bayesian dynamic framework [see Ouysse and Kohn (2009)] and add consumption growth among the economic variables and introduce stochastic idiosyncratic volatility. REIT capitalization has boomed over the last decade and a legislative intervention in 1993 is commonly considered to have made REITs more similar to stocks. Our results do not support this view of a *sudden* change in public real estate pricing characteristics around the early nineties. Yet, idiosyncratic risk shows a marked upward trend with a decreasing rate just in the last years while the quantity of priced market risk has risen since the new century. Our time-varying multi-factor pricing model turns out to be rejected for the asset menu under consideration.

*This is a joint paper with Massimo Guidolin (Manchester Business School, MAGF, and Federal Reserve Bank of St. Louis) and Francesco Ravazzolo (Norges Bank, Research Department). I acknowledge financial support from the Marie Curie Early Stage Training Programme.

1 Introduction

REITS (Real Estate Investment Trusts), estabilished in the United States in 1960, constitutes 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. These specific features have attracted the attention of academics and practitioners on publicly traded real estate as an additional investment opportunity to more traditional options.[1] A popular view, for instance, depicts them as similar to defensive and small cap stocks.

The aim of this paper is, indeed, to analyze the nature and similarity of securitized real estate to other kinds of financial assets within an asset pricing model. Since Chan, Hendershott and Sanders (1990), multibeta models including predetermined macroeconomic factors (inflation, term spread, risk spread, stock capitalization, industrial production among others) have been shown to explain a significant proportion of variation in equity real estate investment trusts (EREIT) returns. Based on previous evidence, Karolyi and Sanders (1998) allow for time-varying risk premiums and betas in a similar framework to Ferson and Harvey (1991) and they find that both stock market and term structure risk premiums are important. Hence, they conclude that REITs are an hybrid of stocks and bonds in terms of their risk exposure and that changes in the prices of economic risks are more important than changes in the betas to explain the predictable variation in REIT returns.

We extend their analysis with recent observations to see whether and how REITs' characteristics have evolved. The market has indeed undergone a remarkable development in the early nineties, the main trigger being the Revenue Reconciliation Act of 1993 aimed at promoting large scale investments of institutional investors. It suffices here to recall that the total market capitalization of the equity REIT sector has increased from about $ 10 bn before the legislative intevention to over $ 300 bn at the end of 2005.[2]

[1] Among others, Fugazza, Guidolin and Nicodano (2009) analyze the ex-post welfare gains of a buy-and-hold strategy when enlarging the asset menu with public real estate and exploiting predictability while Fugazza, Guidolin and Nicodano (2007) focus on the ex-ante benefits from long-run invesing in European real estate.

[2] As an effect of the recent turmoil in the financial markets, the total capitalization at the end of 2008 has fallen to $ 191,651 bn. Also the total number of REITS has decreased, notably by a 30 %. Additional figures are

The recent literature has investigated the consequences of this structural transformation in different directions. For example, Glasgock, Lu and So (2000) argue that the legislative intervention in 1993 in the U.S. may be a watershed that has made securitized real estate more similar to stocks and less to bonds. On the other hand, Clayton and Mackinnon (2003) find that REIT returns have become more linked to unsecuritized real estate.[3]

With respect to Karolyi and Sanders (1998), we add a further risk variable, consumption growth, based on Ling and Naranjo (1996) and introduce some methodological changes. As for the latter, 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 and Kohn (2008). Furthemore, we estimate the sequence of risk premiums following the novel approach by Ouysse and Kohn (2009) to consistently overcome the problems with generated regressors: we show that this approach helps reduce the extent of variations in estimated risk premiums.

In short, we do not find evidence of a sudden shift in public real estate characteristics around the early nineties. Yet, we do find evidence of different kinds of changes: idiosyncratic risk shows a marked upward trend with a decreasing rate just in the last years while market risk appears to have risen starting from the new century and not before, exactly when bonds portfolios comove less with the stock market.

When comparing REITs with other asset classes, we do not observe any strong and convincing similarity with small capitalization stocks. Within the same class of securitized real estate, some heterogeneity is evident between Mortgage and Equity REITs. In more general terms, our multibeta model proves disappointing as the implied mispricing error results significant.

The remainder of the paper is organized as follows. Section 2 outlines the theoretical model we entertain while Section 3 presents our approach. Section 4 describes the variables used in the analysis. Section 5 reports the main results. The concluding section summarizes our findings and outlines further extensions and checks.

---

available from NAREIT (National Association of Real Estate Investment Trusts).

[3] In short, the presence of both a wider analyst following and more sophisticated investors are believed to be the main forces that make prices better reflect property market fundamentals.

2 The Multifactor model

Multifactor models have been used to explain variation in REIT returns. While Titman and Warga (1986) extract the relevant factors, Chan et al. (1990) identify the same factors as in Chen, Roll and Ross (1986): term and default risk premiums, expected inflation, unexpected inflation and changes in industrial production. They find that REITs resemble small capitalization stocks and that the bond market premiums are important to explain the average variation in REIT returns.

The former result is also found by Liu and Mei (1992) who address time variation in risk premiums in a multifactor latent-variable model with four forecasting variables: yield on one-month Treasury bills, the yield spread between AAA bonds and Treasury bills, the dividend yield on an equally weighted stock portfolio and the capitalization rate on equity REITs. The last one proves to be a relevant explanatory variable for both REIT and small cap stocks. Ling and Naranjo (1996) consider time-varying risk premiums with an approach similar to Chan et al. (1990) and find that bond market risk premiums become insignificant after including consumption, whose risk premium instead turns out to be relevant.

Our multibeta model is based on Karolyi and Sanders (1998) and Ferson and Harvey (1991) who relate return predictability to variation in the expected compensation for risk. They posit a linear relationship between returns and a set of macroeconomic factors (namely the stock market return, the default spread, the term spread, unexpected inflation and changes in industrial production) which is subject to variation over time:

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

In the conditional version of Ross'a (1976) APT or Merton's (1973) ICAPM, the expected return on asset i from time $t-1$ to t is related to its factor betas and the associated risk premiums:

$$E(r_{i,t}|Z_{t-1}) = \lambda_0(Z_{t-1}) + \sum_{j=1}^{k} \lambda_j(Z_{t-1})\beta_{ij,t-1} \tag{2}$$

where both the betas and risk premiums are conditional on the information publicly available at time *t-1*, denoted by Z_{t-1}. In the following, we follow Ferson and Harvey (1991) and depart from Karolyi and Sanders (1998) in that our dependent variables are returns in excess of the risk-free rate proxied by the 1-month T-bill.[4] We believe this specification to be more consistent with the mainstream theoretical models.

3 Methodology

Our goal is to investigate the main determinants of REITs returns and consequently trace the differences with respect to competing assets like stocks and bonds by means of the multifactor model presented at the end of the previous section.

Karolyi and Sanders (1998) use a two-stage procedure *à* la Fama and MacBeth (1973) in order to estimate time varying factor sensitivities and risk premiums: the former are obtained through an OLS rolling window over the previous 60 months and then the latter from the cross-section of returns on the estimated betas.

We adopt instead a Bayesian approach and implement some modifications in the model and the estimation procedure. As for the first pass, we specify a state space model of the following form

$$\begin{aligned} r_{i,t} &= \beta_{i0,t} + \sum_{j=1}^{m} \beta_{ij,t} F_{j,t} + \sigma_{it}\epsilon_{it} \\ \beta_{ijt} &= \beta_{ij,t-1} + k_{1ij,t}\eta_{ij,t} \qquad j = 0, ., m, \\ ln(\sigma_{it}^2) &= ln(\sigma_{it-1}^2) + k_{2i,t}\eta_{2i,t} \end{aligned} \tag{3}$$

where $\epsilon_{it} \sim N(0,1)$, $\eta_{i,t} = (\eta_{i0,t}, ..., \eta_{im,t}, \eta_{2i,t}) \sim N(0, Q_i)$ with Q_i a diagonal matrix and

[4] The same transformation applies also to the market return, as customary.

elements $q^2_{i,0}, ..., q^2_{i,m}, q^2_{2i}$. Stochastic variations in betas and log variance 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 $k_{*ij,t}$ indicate the presence of changes in $\beta_{ij,t}$ and $ln(\sigma^2_{it})$ and are not correlated among one another and across time. This specification is very flexible as it allows for both constant and time varying parameters.

We follow Giordani and Kohn (2008) who suggest being quite informative about the size of breaks (Q) for these to be actually found and avoiding diffuse priors on the break frequency as well.[5] The modelling of stochastic betas is believed to be crucial in order to get reliable estimates: as Jostova and Philipov (2005) point out, OLS betas capture the average level of systematic risk but they do not track very well the time pattern especially when betas are very persistent. Indeeed the latter is very likely to be the case from both a theoretical and empirical point of view.[6] Stochastic volatiliy has also become a popular feature in the empirical literature: the presence (and the modelling) of time variation in the conditional second moment is a further element that may render “old style” OLS betas not useful.

In the second pass we estimate the following cross-sections

$$r_{i,t} = \lambda_{0,t} + \sum_{j=1}^{k} \lambda_{j,t}\beta_{ij,t|t-1} + e_{it} \tag{4}$$

where $e_{it} \sim N(0, \sigma^2_{it})$ and $\beta_{ij,t|t-1}$ measures the expected sensitivity of asset i to factor j. The latter is constructed carefully: in our framework, it is obtained by taking the lagged value from the updating step of the Kalman filter and simulating the occurence of future breaks and the shock magnitude from the appropriate posteriors.

The novelty is that the second pass is performed similarly to Ouysse and Kohn (2009) to overcome the notorious error-in-variables problem that plagues traditional studies in small

[5]Details about the algorithm and priors are provided in the Appendix.

[6]Stochastic betas are also crucial ingredients of conditional asset pricing models that have been found to solve typical anomalies associated with unconditional models: the main argument is that the latter neglect the dynamics in (systematic) risk and, hence, are misspecified. 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 an AR(1) process) in the betas. 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. Anyway other empirical studies provide conflicting results and a detailed review is beyond the scope of this paper.

samples. The estimated factor sensitivities from the first pass are likely to be biased and the Bayesian approach provides an elegant way to take into account estimation uncertainty by averaging over parameter draws. In particular, we use the full posterior distribution of the factor sensitivities: for each draw of the betas at a given time t for all the portfolios considered, corresponding values for the risk premia are drawn from the relevant posterior distribution.[7] Then for each time t we end up with an empirical distribution of a given large number of draws of $\lambda_{j,t}$ on which to base our inference on. Here and in the following regressions we make our priors as less infuent as possible in order to let the data speak.

The cross section regression decomposes returns at each time period in a component related to risk, represented by $\sum_{j=1}^{m} \lambda_{j,t}\beta_{ij,t|t-1}$, and a residual $\lambda_{0,t} + e_{it}$. According to any multifactor model, return predictability should be solely due to risk-related components: Karolyi and Sanders (1998) explore the validity of this condition for all the portfolios under analysis and check how much of the predictable return variation is actually captured by the model at hand. In detail, firstly each return is regressed onto a set of instrumental variables that proxy available information at time $t-1$, Z_{t-1}, and the variance of the resulting fitted values computed, $Var(P(r_{it}|Z_{t-1}))$. Then, for each asset i a time series of risk exposures, $\sum_{j=1}^{m} \lambda_{j,t}\beta_{ij,t|t-1}$, is derived and regressed onto the instrumental variables to compute the sample variance of fitted values, $Var(P(\sum_{j=1}^{m} \lambda_{j,t}\beta_{ij,t|t-1})|Z_{t-1})$. Instead the predictable component not captured by the model is the sample variance of the fitted values from the regression of the residuals u_{it} on the instruments, $Var(P(r_{it} - (\sum_{j=1}^{m} \lambda_{j,t}\beta_{ij,t|t-1}))|Z_{t-1})$. The resulting variance ratios

$$\begin{aligned} VR1 &= \frac{Var(P(\sum_{j=1}^{m} \lambda_{j,t}\beta_{ij,t-1})|Z_{t-1})}{Var(P(r_{it}|Z_{t-1}))} \\ VR2 &= \frac{Var(P(r_{it} - \sum_{j=1}^{m} \lambda_{j,t}\beta_{ij,t|t-1}|Z_{t-1}))}{Var(P(r_{it}|Z_{t-1}))} \end{aligned} \tag{5}$$

[7] Details about the exact form of the posterior distributions are provided in the Appendix. An alternative route would be to use mean or median estimates of the betas in a "traditional" regression at each time t: the estimation uncertainty would still be ignored, though. We experimented with this approach as well: see section 5.4.

should equal one and zero respectively if our asset pricing model is correct.[8] We carry out the same test procedure but we are consistent with the approach put forward for the second pass in (4). Basically, we compute VR1 and VR2 for each pair of $\lambda_{j,t}$ and $\beta_{ij,t|t-1}$ from their posterior distribution: consequently, we have an empirical distribution for our quantities of interest.

Finally, the predictable variation of returns due to the multiple-beta model is decomposed into the components due to each economic variable: we compute $Var(P(\lambda_{j,t}\beta_{ij,t|t-1})|Z_{t-1})$ for each factor k and relate it to $Var(P(\sum_{j=1}^{m}\lambda_{j,t}\beta_{ij,t|t-1})|Z_{t-1})$. The sum of these components should not equal that of the combination because of the covariance among risk prices and betas.

4 Data Description

The variables used in our analysis are listed in Table 1: they belong to three main categories. The first one, Portfolio Returns, includes several asset classes like stocks and bonds. The Stocks are publicly traded firms listed on the NYSE, AMEX and Nasdaq which are sorted into industry and decile portfolios by their four-digit SIC code and market capitalization, respectively. The bonds are high-yield corporate bonds, intermediate-term and long-term government bonds. REIT return indices are tax-qualified equity REITs.

The economic variables, instead, are the standard proxies for the risk factors potentially priced in asset returns like, for example, the excess market portfolio return and the default premium. Here we add consumption growth to the set of variables considered by Karolyi and Sanders (1998). Finally, the instrumental variables approximately represent the information set investors have when making portfolio decisions.

We use monthly observations for a sample ranging from 1983:01 to 2008:07. Actually we estimate the model over a longer sample starting in 1979:12 but we restrict the analysis to the shorter one. The reason is twofold: firstly, we want to represent a situation in which investors start with a weak prior belief of return unpredictability but have also some observations available to learn from; secondly, we want to be consistent with Karolyi and Sanders (1998) as for the

[8] 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.

initial date.

5 Results

5.1 Descriptive statistics

A quick glance at the first two sample moments in Table 2 suggests interesting considerations: first, we do not observe any size premium unlike conventional beliefs while sharp differences within the REITs' category are evident. In particular, NAREIT Mortgage and NAREIT Hybrid (NARMTG and NARHYB, respectively, in Table 1) have almost a null excess return along with a standard deviation sensibly higher than other REITs. Equity REITs are usually depicted as defensive stock: they indeed have first and second sample moments almost identical to Utilities. In short, this superficial comparisons reinforce the common belief that securitized real estate is a hybrid of stocks and bonds while it cannot tell us much about changes in characteristics over the last few years. Looking at all the columns of the same Table, we can only notice that the market for public real estate has undergone a radical change in the early nineties that has manifested itself in first moments while second moments have jumped upward more recently.

5.2 Idiosyncratic Risk

We firstly comment on the path of the residual variance for each group of assets from the estimation output of (4): Figures 1-3 report the median of the posterior distribution of each σ^2_{it} together with its 50% credibility interval.[9]

As for REITs, we basically observe in Figure 3 a slight upward trend that seems to stabilize in the last part of the sample, more precisely starting from the second half of 1998: there is no evidence of one-time shift but instead of a continuous increase in idyosincratic risk which reaches levels more similar to those of small stocks rather than bonds.[10] Mortage and Hybrid

[9] As already specified in the previous section, we opted for a tight prior on the break occurence which favours rare changes. Inference is anyway unaffected in this case by using looser priors. The influence of the initial conditions completely disappears as a signal of convergence of the algorithm.

[10] Even though our priors are very supportive of rare changes of small magnitude, note that the algorithm is flexible enough to accomodate both continuous and discrete shifts. Furthermore, the use of smoothing algorithms make final estimates much less jagged as evident in the graphs.

(investing both in equity and mortgages) REITs, the most volatile in our sample, show also a bounce in the middle of 2007 which can be thought of as coincident with the onset of the current financial crisis. Ooi, Wang and Webb (2009) compute measures of idiosyncratic volatility of the average REIT stock from 1990 to 2005 using the three-factor model of Fama and French (1993) and find that it displays an asymmetric (counter)cyclical behaviour as idiosyncratic volatility slightly decreases in good times but it readily jumps in bad times. More importantly, the authors note a downward trend that can be explained in light of the impressive increase in the average size of REITs after 1990. As remarked above, we get an exactly opposite conclusion. Even though our approaches are similar up to this stage, they nevertheless differ in some respects that are potentially crucial: the pricing model (we use a multifactor model not including SMB and HML), the time variation in both systematic and idiosyncratic risk (we impose a special kind of persistent process while they leave it unmodelled), the sample (we start from 1983) and the focus (sector-level versus firm-level).[11] On the other hand, we both conclude that the variance of returns is strongly dominated by idiosyncratic risk.

A different picture emerges for the size decile and most of the industry portfolios (see Figures 1-2): they all experience an increase in the residual volatility around 2000 which is then reabsorbed in two-three years. The very few exceptions are the energy industry whose idiosyncratic risk does not seem to fall and the telecom industry that instead does not experience any substantial increase over the sample.[12]

An interesting result stands out if we look within the category of size-sorted portfolios as the differences in the amount of (total) volatility are magnified in the conditional estimates: our

[11] We conjecture that the different stucture we impose in the time varying models may be the relevant aspect but we leave a careful examination for future drafts. As a robustness check we have estimated also a GARCH(1,1) model for each asset and our result of an upward is strongly confirmed. Moreover, Ooi *et al.* (2009) find idiosyncratic volatility to be actually priced in the cross-section of REITs' returns, even after including typical anomalies like size and momentum effects.

[12] Our evidence is confirmed by a handful of recent papers investigating the supposed "idiosyncratic volatility puzzle" pointed by Campbell, Lettau, Malkiel and Xu (2001). They report a positive trend in idiosyncratic volatility during the 1962 and the 1997 period while the aggregate market and industry volatilities remained roughly constant. Brandt, Brav, Graham and Kumar (2008) and Zhang (2008) note that the trend has strongly reverted from 2001 onwards to start climbing up again since 2007: it means that the long-lasting rising volatility was merely an episode that calls for a sound economic explanation. The actual break is located by Brandt, Brav, Graham and Kumar (2008) around mid-late 2000 as it also appears in our graphs, while Zhang (2008) reports very similar results for the same sectors we consider. All in all, it is reassuring that different, even though not contrasting, approaches deliver the same message.

macroeconomic factors seem to explain considerably well the largest cap portfolios.

It has to be borne in mind that the focus in this section is just on unexplained variance and not total variance: our measures of idiosyncratic risk intrinsecally depend on the set of chosen factors. The path we find for public real estate can also be interpreted as a signal of worsened fit of our regression model or emergence of new relevant omitted economic variables.

5.3 Factor sensitivities

For our purposes, it is interesting to examine the pattern of estimated betas in (4). Figures 4-11 report the sequence of median β_{it}'s for each economic variable but we focus mainly on factor sensitivities for REITs.[13]

What we call here β_0, usually referred to as α, represents a measure of the risk-adjusted excess return implied by the model at hand: therefore an accurate investigation of its dynamic properties is of primary interest.[14] It has predominantly a negative sign and varies smoothly but zero is always included within the 90% posterior bands, particularly after 2000.[15] It also displays a gentle decline in the last two years approximately.

The sensitivity to the stock market return factor (VW) is estimated very precisely. We notice that REITs show an unpward jump around 2000. Size decile portfolios exhibit a beta about unity while REITs' one wander close to one half. On the contrary, government bonds display a downward trend leading the coefficient very close to zero.

The sensitivity to the default premium (PREM) for securitized real estate has a positive sign and shows a cyclical pattern with a recent declining trend after a bounce at the outset of the new millenium, spiking in 2001: only in the last part of our sample (after 2000) the 90% credible interval contains zero. Small cap portfolios offer a similar picture even though in a less

[13]The graphs also include the 90% Bayesian credible intervals. As we want to draw comparisons, we collect the plots of all the portfolio betas for each economic variable.

[14]Our specification follows the prevalent approach of considering excess returns on the left hand side and, consequently, an immediate theoretical restriction to test is whether β_0 is null on average and/or at each time t. Karolyi and Sanders (1998), on the other hand, opt for raw realized returns and cannot provide a plain economic interpretation for the intercept.

[15]Actually, for mortgage REITs the zero is most of the time outside our posterior bands and there is evidence of negative risk-adjusted excess returns. Besides, mortgage and hybrid REITs exhibit greater values for the intercept: this could be taken as a signal that our multifactor model does not correctly price such subcategories of public real estate.

sharp way.

As for the relation between REITs returns and changes in the term structure (DSLOPE), we observe that the degree of uncertainty increases sistematically in the late nineties. The sign of the coefficient is positive for Mortgage REITs and negative for Equity and Composite REITs: intuitively, mortgages are tied up with long term rates which on the other hand affects also the value of long term lease contracts, this time negatively through discounting.

REIT returns have a negative linear relationship with changes in the level of economic activity (IPGRW) in a similar way to small cap stocks and some industry portfolios (namely, Utilities).

Finally, for unexpected inflation (UI), consumption growth (CONSUM) and the real interest rate factor (REALTB) the sentivities do not appear to be significant once we take into account measures of parameter and estimation uncertainty.

We can conclude that the most interesting findings here is that real estate has been experiencing an increasing comovement with the stock market return as a whole in the last years and that there exists some sort of heterogeneity between the different REIT portfolios.

5.4 Risk premiums

Once the first pass is run as in (3), we estimate the risk premiums from (4) following the approach outlined in section 3. The estimates are evidently characterized by a great deal of uncertainty: the bands tend to widen simultaneously in some specific moments such as the second half of 1987 and the period between the late nineties and the beginning of the twenty-first century.[16]

It is interesting to analyze how the (median) $\lambda_{j,t}$ are distributed over time. Table 3 reports means, 5th and 95th percentiles for the full sample and two subsamples (1983:01-1992:12 and 1998:01-2007:12). The amount of variation in risk premiums is quite evident from this table as well: empirical distributions are very dispersed, especially for the market risk premium (VW). When looking at subsamples, premiums are found to be more volatile in the last ten years (1998:01-2007:12): the neater example is PREM which exhibits severe fluctuations during the

[16]The 90% posterior credible sets basically always contain zero except for the price of market risk which appears to be estimated more precisely. In order to save space we do not report graphs in this draft but they are available upon request. Anyway, it has to be borne in mind that estimates of risk premiums are well known to be very noisy and subject to some drawbacks.

first years of the new century.

To test for the significance of the average risk premiums, numerical standard errors are reported.[17] In the full sample, only ZBETA (the intercept or also λ_0 as we refer to in the formulas) results significant with CONSUM as a borderline case. Breaking down the entire sample provides useful insights: DSLOPE and REALTB are significant until 1992:12. The stock market risk premium (VW) turns out not to be priced on average all over the sample and particularly in the second subsample.[18,19]

As a further check, a Bayesian variant of the Fama-MacBeth's statistic is also computed. In particular, for each draw i of the risk premiums ($\lambda_{j,t}$) we compute the average over time ($\widehat{\lambda}_j^i$) in order to finally obtain an empirical distribution of our quantity of interest. Table (4) confirms previous findings for the full sample while DSLOPE e REALTB are not significant anymore in the second subsample considering the 5th and the 95th percentiles. According to the "Bayesian" statistic, at least 97.5% percent of the posterior distribution of the stock market risk factor has positive values in the period 1983:01-1992:12 whereas in the full sample this probability is greater than 95% but less than 97.5%. Anyway the two metrics do convey an important message: there exists a non-zero pricing error (captured by ZBETA). Interestingly, it emerges only after the early nineties and is more precisely estimated since then until 1997 if we compare the percentile in each section of Table (4).

As already mentioned in section 3, the second pass has been performed following Ouysse

[17] The t-ratios la Famma-MacBeth (1973) assume the absence of autocorrelation over time in the (median) risk premiums while it is mildy present, even though only in the very first lags actually. Regardless of the way we compute standard errors, note that this kind of inference implicitly embraces a frequentist perspective.

[18] Ooi *et al.* (2009) estimate for each month a cross-section with explanatory variables like beta, size, book-to-market equity ratio, past return and idiosyncratic risk for only REITs. They similarly find that the market beta is not priced in their sample and their result is robust to several specifications. More importantly, when the market beta is the only independent variable (a model not too dissimilar from ours given that the other risk factors have a negligible impact on the test asset returns) the average intercept is strongly significant. We note anyway that our analysis has a larger scope and looks at other assets as well.

[19] As many studies in empirical finance fail to find a systematic relationship between returns and market betas, part of the recent literature advocates conditional approaches in which the researcher distinguishes across market conditions. The main argument is that, when using realized rather than expected returns, negative premium estimates are very likely to be driven by market crashes. See Koch and Westheide (2008) and references therein for more details. Related to our focus, Conover et al. (2000) find that the lack of significance of the market risk premium for equity REITs vanishes when using a dual-beta model. In bull months, the beta is positive and significantly related to EREIT returns: it implies that at good times, the riskier (in terms of beta) a stock is the higher its expected return.

and Kohn (2009) in order to avoid what frequentists call the generated-regressor problem.[20] We also tried the simpler and traditional approach of using plain (median, in our case) betas from the first pass rather than the full posterior distribution at each time t.[21] While there are no sharp differences in terms of uncertainty sorrounding each λ_j at time t, the picture is somewhat different for the empirical distribution of the $\widehat{\lambda}_j$'s.[22] The increased dispersion in the estimates is striking: in particular, the extent of fluctuations in PREM and RealTB is impressive compared to Table 3. Given that such degrees of variation are higly suspect, we prefer Ouysse and Kohn (2009) when balacing complexity and reliability.

5.5 Decomposing predictable variation

The multifactor model we employ states that the predictable variation in asset returns should be fully captured by the component related to risk. Information publicly available at each time t is proxied by a set of instrumental variables listed in Table 1 plus a dummy variable for the month of January to account for the January effect. Following Karolyi and Sanders (1998), we compute the two statistics in (5): anyway, we implement the same approach that informs the second pass estimation, as explained in section 3, and finally normalize them to one.

Results in Table 6 confirm our previous evidence that the multifactor model falls short of capturing adequately the variation in expected returns. Anyway, our approach (consistent with Ferson and Harvey (1991)) is doomed to deliver less compelling results that Karolyi and Sanders' as they treat λ_0, a zero-beta portfolio which turns out to play on average an important role in the cross-section of returns, as a risk-related component that enters VR1. In our framework, instead, it has to belong to the residual part: in fact our values of VR2 are always high and sometimes comparable to VR1's.

Among the CRSP size portfolios, the model works better for medium and large cap portfolios. Real estate portfolios have values for VR1 which are very similar to those of small cap stocks

[20]Alternatively, from a Bayesian point of view the factor betas are random variable themselves and we would ignore estimation uncertainty otherwise.

[21]To save space we skip graphs while Table 5 reports the main results.

[22]We note that closed form solutions are available for $\lambda_{j,t}$ with this approach: hence, posterior credible intervals are easily obtained. Instead, with the Ouysse and Kohn (2009) approach, simulations are needed.

and some industry portfolios but higher than those of government bonds.

The last step consists of breaking down the model-implied risk related predictable variation of returns into its economic constituents. We unsurprinsingly find that he biggest role for stocks is played by the stock market portfolio return while the overall picture is more balanced for REITs as also the default premium (PREM) contributes in an important way. The considerations in section 5.3 apply here as well: in particular, note the different importance of industrial production growth for mortgage and equity REITs.

6 Robustness checks

6.1 Risk premiums without REITs

In addition we have repeated the same computations as in section 5.4 in absence of real estate returns to verify whether their inclusion is at the heart of our model empirical rejection (see Table 7). A couple of interesting considerations are at effect: 1) the degree of uncertainty (in other terms, the dispersion of the posterior distribution) is smaller now; 2) more importantly, in the full sample VW also becomes significant even though as a borderline case considering the numerical standard error and most of the large (in absolute terms) estimates have now disappeared.[23]

According to the Bayesian version of the Fama-MacBeth's statistics (see Table 8), as before there is more than 95% but less than 97.5% probability that the market risk premium is positive in the full sample. The same holds for the first subsample while before the evidence was stronger. As for the intercept (ZBETA), it still remains significant in the long sample while this is not anymore the case in both subsamples according to convenional levels of uncertainty (confidence or credibility in classical or Bayesian jargon, respectively).

To sum up, even when we restrict the attention to stocks and bonds, our multifactor model is not supported by the data that instead seem to favour the simpler CAPM-style specification.[24]

[23]Further findings are: a) consumption growth is significant only in the first subsample and displays a negative sign; 4) finally, in the second subsample only PREM and DSLOPE (the latter with a negative sign) turns out to be significant while in the complete specification this was the case just for consumption growth.

[24]Properly speaking, even a CAPM-related one factor model is going to be to be rejected by the data as our

The only notable difference is that now the strength of evidence for significant mispricing (or absence of relevant factors) is weaker: this is clearer in the 1998-2008 period.

6.2 Informative priors

Finally we have experimented with an informative prior in the second pass in order to put some structure (constraints) in the distribution and moments of the risk premiums. These are postulated to be normally distributed with zero mean and variance such that there is 95% probability that annualized premia are in absolute term smaller than the maximum return observed in the sample for all the assets.[25]

A striking finding is the remarkable reduction in the variability of the estimated *premia* with respect to the baseline case.[26] Interestingly, the stock market risk premium is significant in the whole and first sample. So, this further check confirms our conclusions at the end of the previous subsection (6.1).

7 Conclusions

In this work we aim to analyze and compare publicly traded real estate with general asset classes like stocks and bonds by using a multifactor pricing model inspired by Karolyi and Sanders (1998). Yet, we do not focus only on the "relative" degree of predictability but we also examine other aspects of securitized real estate. As in most of the recent empirical literature, we model the evolution of both sensitivities to risk factors and idiosyncratic risk as latent stochastic variables: in particular, we use the mixture innovation approach, proposed by Ravazzolo et al. (2007) and Giordani-Kohn (2008), capable to capture structural shifts. We prefer such a feature as capitalization has boomed over the last decade and a legislative intervention in 1993 is considered to have made REITs more similar to stocks.

Our results do not support the idea of a sudden change in public real estate characteristics around the early nineties. On the contrary, idiosyncratic risk shows a marked upward trend

estimates of λ_0 seem to suggest.

[25] A complete description of prior distributions and hyperparameters used are in the Appendix.

[26] For the sake of brevity, we do not report tables but they are available from the authors.

with a decreasing rate just in the last years and market risk appears to have risen only since the new century when also bonds portfolios comove less with the stock market. When comparing REITs with other asset classes, we do not observe any strong similarity with small capitalization stocks. Within the same class of securitized real estate, some heterogeneity is evident between Mortgage and Equity REITs.

There is evidence of mispricing, though: our measure (λ_0) results significant regardless of the metric or test we use. A natural check is to leave out public real estate and assess the validity of our model for stocks and bonds only. A preliminary analysis suggests that the multibeta specification does not even pass this test. Recent works in the empirical asset pricing literature points to an important role played by idiosyncratic risk, usually neglected. On the other hand, some important economic risk variable may be missing. However, we note that the bulk of evidence for such a mispricing dates back to the early nineties and not in the very last part of the sample.

On the methodological side, we apply the novel approach of Ouysse and Kohn (2009) for the estimation of risk premiums in a coherent Bayesian framework. Integrating the uncertainty in the estimation of the factor sensitivities seems to be a sensible strategy in the present context.

References

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

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

[3] 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.

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

[5] Chan K. C., Henderhott P. and Sanders A. B., “Risk and Return on Real Estate: Evidence from Equity REITS”, *AREUEA Journal*, 1990, 18, 4, 431452.

[6] Chui A.C.W., Titman S. and Wei K.C.J., “The cross section of expected REIT returns”, *Real Estate Economics*, 2003, 31.

[7] Clayton J. and MacKinnon G., “The Relative Importance of Stock, Bond and Real Estate Factors in Explaining REIT Returns”, *Journal of Real Estate Finance and Economics*, 2003, 27, 1, 39-60.

[8] Conover M., Friday H. and Howton S., “An Analysis of the Cross Section of REturns for EREITs Using a Varying-Risk Beta Model”, *Real Estate Economics*, 2000, 28, 1, 141-163.

[9] Fama E. and French K.R., “Common Risk Factors in the Returns on Stocks and Bonds”, *Journal of Financial Economics*, 1973, 81, 607636.

[10] Fama E. and MacBeth J., “Risk, Return and Equilibrium: Empirical Tests”, *Journal of Political Economy*, 1973, 81, 607636.

[11] Ferson W. and Harvey C. R., “The Variation of Economic Risk Premiums”, *Journal of Political Economy*, 1991, 99, 385415.

[12] Fugazza C., Guidolin M. and Nicodano G., “Investing for the Long-run in European Real Estate”, *Journal of Real Estate Finance and Economics*, 2007, 34, 35-80.

[13] Fugazza C., Guidolin M. and Nicodano G., “Time and Risk Diversification in Real Estate Investments: Assessing the Ex Post Economic Value”, *Federal Reserve of St. Louis Working Paper*, 2009, N. 001A.

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

[15] Giordani P. and Kohn R., “Efficient Bayesian Inference for Multiple Change-Point and Mixture Innovation Models”, *Journal of Business and Economic Statistics*, 2008, N.26 (1), 66-77.

[16] Glascock J., Lu C. and So R., “Further Evidence on the Integration of REIT, Bond and Stock Returns”, *Journal of Real Estate Finance and Economics*, 2000, 20, 2, 177-194.

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

[18] 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.

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

[20] Ling, D.C. and Naranjo A., “Economic Risk Factors and Commercial Real Estate Returns”, *Journal of Real Estate Finance and Economics*, 1997, 15, 3, 283307.

[21] 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.

[22] Merton R., “An Intertemporal Capital Asset Pricing Model”, *Econometrica*, 1973, 41, 867-887.

[23] Ooi J.T., Wang J. and Webb J., “Idiosyncratic Risk and REIT Returns”, *Journal of Real Estate Finance and Economics*, 2009, 38, 420-442.

[24] Ouysse R. and Kohn R., “Bayesian Variable Selection and Estimation of Risk Premiums in the APT model”, *Discussion Papers*, School of Economics, The University of New South Wales, 2009.

[25] Pesaran M. H. and Timmermann A., “Predictability of Stock Returns: Robustness and Economic Significance”, *Journal of Finance*, 1995, 50, 1201-1228.

[26] 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.

[27] Zhang C., "A Re-examination of the Causes of Time-varying Stock REturn Volatilities", *Working Paper* , 2008.

8 Appendix

As specified in section 3 we use a Bayesian variant of the two step procedure à la Fama - MacBeth (1973). Here we present each step in detail.

8.1 First pass

For each asset, the model in (1) is

$$\begin{aligned} r_t &= \beta_{0,t} + \sum_{j=1}^{m} \beta_{ij,t} F_{j,t} + \sigma_{it}\epsilon_{it} \\ \beta_{ijt} &= \beta_{ij,t-1} + k_{ij,t}\eta_{ij,t} \qquad j = 0, ., n, \\ ln(\sigma^2_{it}) &= ln(\sigma^2_{it-1}) + k_{2i,t}\eta_{2i,t} \end{aligned}$$

where $\epsilon_{it} \sim N(0,1)$, $\eta_{i,t} = (\eta_{i0,t}, ..., \eta_{im,t}, \eta_{2i,t}) \sim N(0, Q_i)$ with Q_i a diagonal matrix and elements $q^2_{i0}, ..., q^2_{im}, q^2_{2i}$, and $k_{it} = (k_{i0,t}, ..., k_{im,t}, k_{2i,t})'$ is a $((m+2) \times 1)$ vector of unobserved uncorrelated 0/1 processes with $\Pr[k_{ijt} = 1] = \pi_{ij}$ for $j = 0, ., m+1$.

The model parameters are the structural break probabilities $\pi_i = (\pi_{i0}, ..., \pi_{im}, \pi_{2i})'$ and the vector of variances of the break magnitude $q^2_i = (q^2_{i0}, ..., q^2_{im}, q^2_{2i})$. They are collected in a $(2(m+1) \times 1)$ vector $\theta_i = (\pi_{i0}, ..., \pi_{im}, \pi_{2i}, q^2_{i0}, ..., q^2_{im}, q^2_{2i})'$.

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

$$\pi_{ij} \sim Beta(a_j, b_j) \tag{6}$$

where the hyperparameters a_j and b_j reflect the prior belief about the occurence of breaks. For the variance parameters the inverted Gamma-2 prior is chosen

$$q^2_{ij} \sim IG(\nu_j, \delta_j) \tag{7}$$

where ν_j expresses the strenght of our 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 $B = \{\beta_t\}_{t=1}^T$, $R = \left\{\sigma_t^2\right\}_{t=1}^T$ and $K = \{k_t\}_{t=1}^T$ are simulated alongside the model parameters θ.

The complete data likelihood function is given by

$$\begin{aligned} p(r, B, K, R|\theta, F) &= \prod_{t=1}^{T} p(r_t|F_t, \beta_t, \sigma_t^2) \prod_{j=0}^{m} p(\beta_{jt}|\beta_{jt-1}, k_{jt}, q_j^2) \times \\ & \quad p(\sigma_t^2|\sigma_{t-1}^2, k_{2t}, q_2^2) \times \\ & \quad \prod_{j=0}^{k} \pi_j^{k_{jt}}(1-\pi_j)^{1-k_{jt}} \pi_2^{k_{2t}}(1-\pi_2)^{1-k_{2t}} \end{aligned} \tag{8}$$

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

$$p(\theta, B, K, R|r, F) \propto p(\theta) p(r, B, K, R|\theta, F) \tag{9}$$

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

1. Draw K_β conditional on R, K_σ, θ, and r.

2. Draw B conditional on R, K, θ and r.

3. Draw K_σ conditional on B, K_β, θ, and r.

4. Draw R conditional on B, K, θ and r.

5. Draw θ conditional on B, K and r.

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

$$\begin{aligned} p(k_{\beta,t}|K_{\beta,-t}, K_\sigma, R, \theta, r) &\propto p(r|K_\beta, K_\sigma, R, \theta)p(k_t|K_{\beta,-t}, \theta) \\ &\propto p(r^{t+1,T}|r^{1,t}, K, R, \theta) \\ &\quad p(r_t|r_{1,t-1}, k_{\beta,1,t-1}, R, \theta, x)p(k_{\beta,t}|K_{\beta,-t}, \theta). \end{aligned} \tag{10}$$

Gerlach *et alia* (2000) show how to evaluate the first two terms while the last one is obtained from the prior. When $K_{\beta,t}$ and β_{jt} are higly 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).

K_σ and R are drawn in the same way as K_β 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 Gibb 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 θ is easily sampled as we use conjugate priors.

We use a burn-in period of 1000 and draw 5000 observations storing every other of them to simulate the posterior distributions of parameters and latent variables. The resulting autocorrelations of the draw are very low.

8.2 Second pass

To estimate the cross section in (4) 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 $(m \times 1)$ vector and N is the total number of assets, we use natural conjugate priors. In particular,

$$p(\lambda, \sigma^2) = p(\lambda|\sigma^2) * p(\sigma^2) \tag{11}$$

where

$$(\lambda|\sigma^2) \sim N(\underline{\lambda}, \sigma^2 \underline{V}) \tag{12}$$

and

$$(\sigma^2) \sim IG(\frac{\underline{\nu}}{2}\underline{s}^2, \frac{1}{2}\underline{\nu}) \tag{13}$$

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

$$\begin{aligned} (\lambda|r) &\sim t(\overline{\lambda}, \overline{s}^2\overline{V}, \overline{\nu}) & (14) \\ (\sigma^2|r) &\sim IG(\frac{\overline{\nu}}{2}\overline{s}^2, \frac{1}{2}\overline{\nu}) & (15) \end{aligned}$$

with

$$\begin{aligned} E(\lambda|r) &= \overline{\lambda} & (16) \\ var(\lambda|r) &= \frac{\overline{\nu s}^2}{\overline{\nu}-2}\overline{V} & (17) \\ E(\sigma^2|r) &= \frac{\overline{\nu s}^2}{\overline{\nu}-2} & (18) \\ var(\sigma^2|r) &= \frac{(\overline{\nu s}^2)^2}{(\overline{\nu}-2)^2(\frac{\overline{\nu}}{2}-2)} & (19) \end{aligned}$$

where

$$\begin{aligned} \overline{V} &= (\underline{V}^{-1} + (X'X)^{-1})^{-1} & (20) \\ \overline{\lambda} &= (\underline{V}^{-1} + (X'X)^{-1})^{-1}(\underline{V}^{-1}\underline{\lambda} + (X'X)^{-1}\hat{\lambda}) & (21) \\ \overline{\nu} &= \underline{\nu} + N & (22) \end{aligned}$$

and $\hat{\lambda}$ is the OLS estimate.[27]

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

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

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.

Table 1: **Economic and Instrumental Variables**

PORTFOLIO	*DEFINITION*
	Portfolio Returns
$NoDur$	Food, Tobacco, Textiles, Apparel, Leather, Toys
$Durbl$	Cars, TV's, Furniture, Household Appliances
$Manuf$	Machinery, Trucks, Planes, Chemicals, Off Furn, Paper, Com Printing
$Enrgy$	Oil, Gas, and Coal Extraction and Products
$HiTec$	Computers, Software, and Electronic Equipment
$Telcm$	Telephone and Television Transmission
$Shops$	Wholesale, Retail, and Some Services (Laundries, Repair Shops)
$Hlth$	Healthcare, Medical Equipment, and Drugs
$Utils$	Utilities
$Other$	Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment, Finance
$Q1 - Q10Size$ $DecilePort$	Ten value-weighted portfolios of NYSE, Amex and Nasdaq stocks
LTG	Long-term government bonds
$BAAC$	Corporate bonds (Baa quality rating)
ITG	Intermediate-term government bonds
$NAREIT$	National Association of Real Estate Equity Investment Trusts, all securities
$NAREQ$	NAREIT-NYSE, Amex and NMS tax qualified equity trusts only
$NARMTG$	NAREIT-Mortgages only
$NARHYB$	NAREIT-Hybrid equity/mortgage trusts
$DJWILREIT$	Dow Jones Wilshire US REIT Total Return Index
	Economic Variables
VW	Value-weighted NYSE, Amex and Nasdaq return less 1m Tbill return
$PREM$	Monthly yield of Baa corporate bonds less long govt bonds
$DSLOPE$	Change in the difference between monthly yield of interm-govt bonds and 1m Tbill yield
UI	Unexpected inflation rate from an ARIMA (0,1,1) model of CPI inflation, s.unad.
$IPGRW$	Monthly growth rate of s.a., real, industrial production
$CONSUM$	Monthly growth rate of total, s.a., real, personal consumption expenditure index
$REALTB$	One month Tbill return less CPI infl.
	Instrumental Variables
$VW(-1)$	Value-weighted NYSE, Amex and Nasdaq return less 1m Tbill return
$HBIT(-1)$	Interm govt bond less 1m Tbill return
$JUNK(-1)$	Yield spread Baa-Aaa
$DIV(-1)$	Monthly dividend yield on S&P's 500 index
$TBILL(-1)$	1 Month Tbill return

Table 2: Descriptive Statistics for portfolio returns, economic variables and instrumental variables in three different time windows.

	1983:01-1992:12		**1983:01-1997:12**		**1983:01-2008:07**	
	Mean	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
NoDur	1.33	5.05	1.22	4.56	0.83	4.24
Durbl	0.62	6.22	0.72	5.53	0.42	5.96
Manuf	0.75	5.27	0.91	4.68	0.76	4.61
Enrgy	0.70	4.90	0.82	4.46	0.90	5.10
HiTec	0.16	6.09	0.69	5.91	0.58	7.28
Telcm	1.01	4.72	1.03	4.39	0.53	5.16
Shops	1.06	5.99	0.90	5.32	0.66	5.11
Hlth	0.97	5.46	1.07	5.13	0.73	4.75
Utils	0.76	3.48	0.73	3.38	0.64	3.91
Other	0.75	5.15	0.97	4.67	0.61	4.77
CAP1RET	0.21	5.51	0.48	5.05	0.62	5.79
CAP2RET	0.44	5.65	0.62	5.15	0.65	6.03
CAP3RET	0.67	5.53	0.78	5.11	0.71	5.59
CAP4RET	0.66	5.58	0.76	5.07	0.64	5.45
CAP5RET	0.79	5.41	0.87	4.91	0.73	5.30
CAP6RET	0.78	5.12	0.85	4.63	0.68	4.88
CAP7RET	0.77	5.04	0.89	4.50	0.76	4.74
CAP8RET	0.77	5.01	0.82	4.45	0.70	4.81
CAP9RET	0.82	4.82	0.92	4.31	0.74	4.34
CAP10RET	0.80	4.41	0.95	4.03	0.59	4.21
LTGOVB	0.29	1.31	0.27	1.22	0.22	1.16
BAACORPB	0.58	1.58	0.52	1.56	0.41	1.57
INTGOVB	0.23	0.83	0.20	0.77	0.17	0.73
NAREIT	0.16	3.12	0.46	3.12	0.48	3.70
NAREQ	0.55	3.33	0.73	3.31	0.67	3.80
NARMTG	-0.36	3.54	0.11	3.85	0.04	5.61
NARHYB	-0.04	3.91	0.30	3.60	0.03	5.25
DJWILREIT	0.27	3.39	0.51	3.46	0.57	3.94
VW	0.71	4.53	0.85	4.07	0.60	4.25
PREM	2.06	0.41	1.93	0.41	2.09	0.52
DSLOPE	0.01	0.78	-0.01	0.71	-0.00	0.62
UI	0.00	0.24	0.01	0.22	0.01	0.29
IP	0.26	0.58	0.32	0.54	0.25	0.53
CONSUM	0.29	0.15	0.25	0.15	0.23	0.17
REALTB	0.25	0.24	0.22	0.23	0.14	0.31
XVW	0.71	4.53	0.85	4.07	0.60	4.25
HBIT	2.03	1.06	1.90	1.02	1.53	1.09
JUNK	0.95	0.37	0.70	0.48	0.55	0.50
DIV	3.63	0.58	3.23	0.77	2.55	1.02
TBILL	6.78	1.82	6.00	1.95	4.94	2.24

Notes: The definition of the portfolio returns, economic variables and instrumental variables listed above is in Table 1.

Figure 1: Conditional variance of the ten industry, the long and the intermediate government bond portfolios.

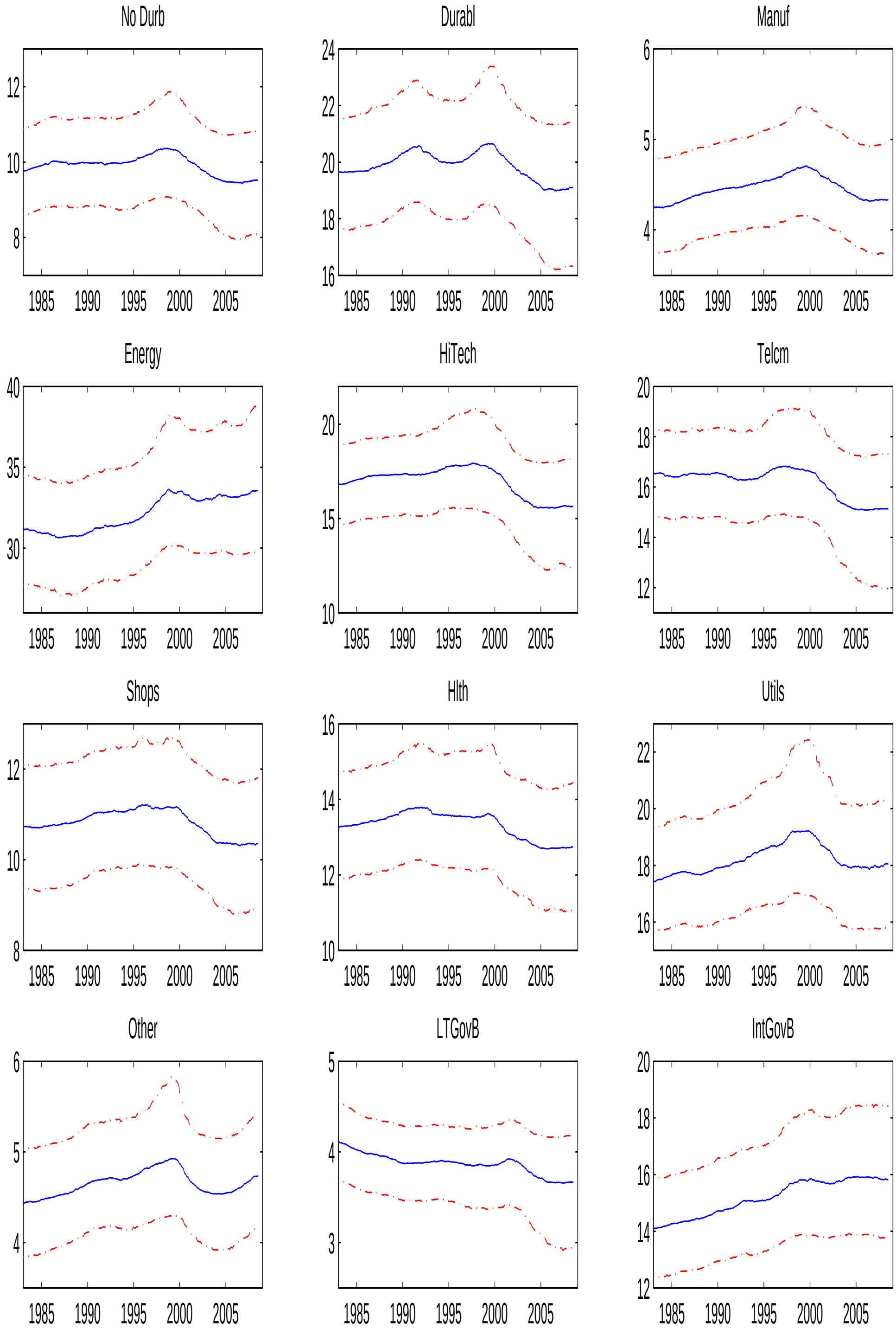

Figure 2: Conditional variance for the ten decile capitalization and the High-Yield corporate bond portfolios.

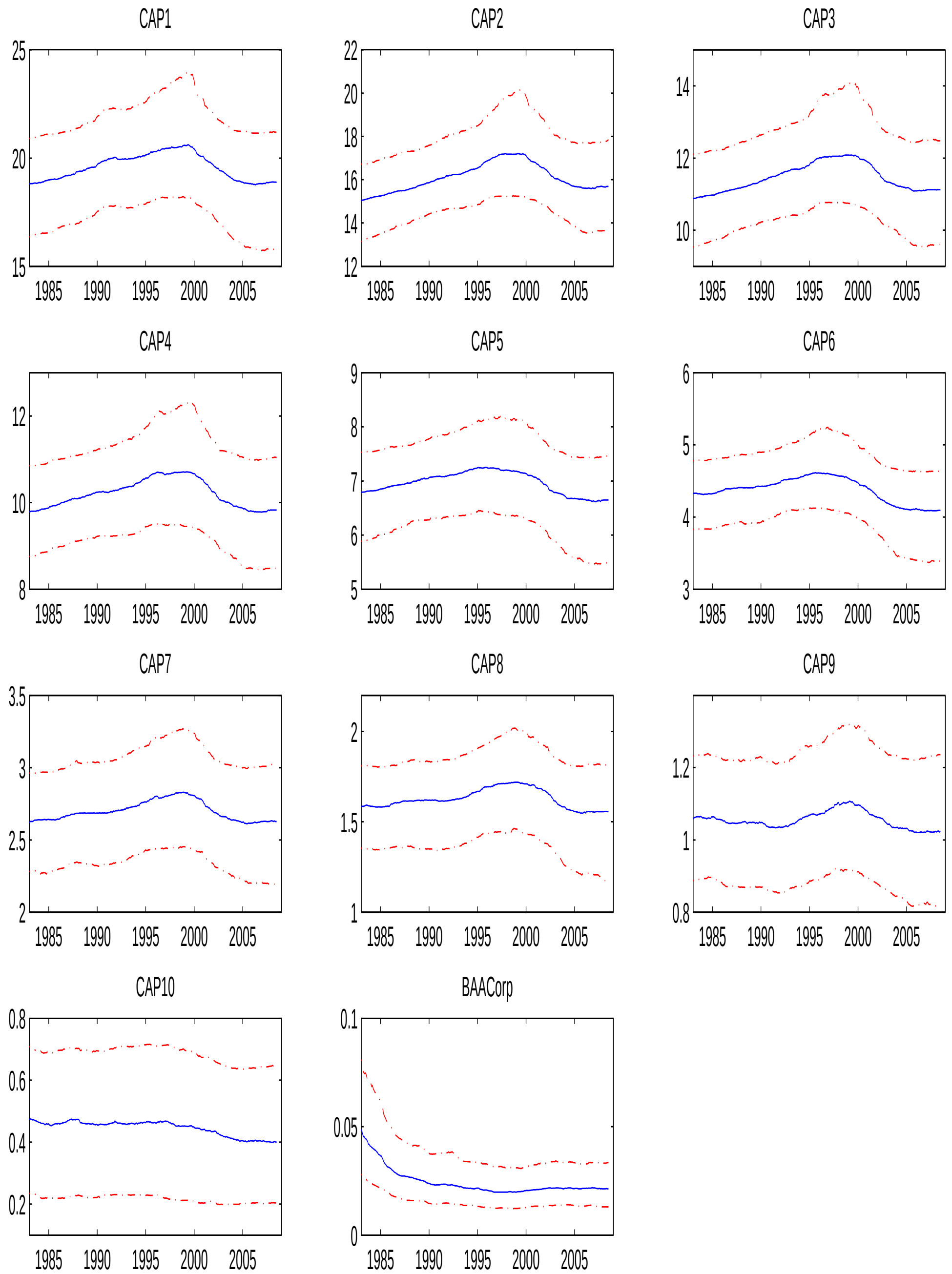

Figure 3: Conditional variance for the five real estate portfolios.

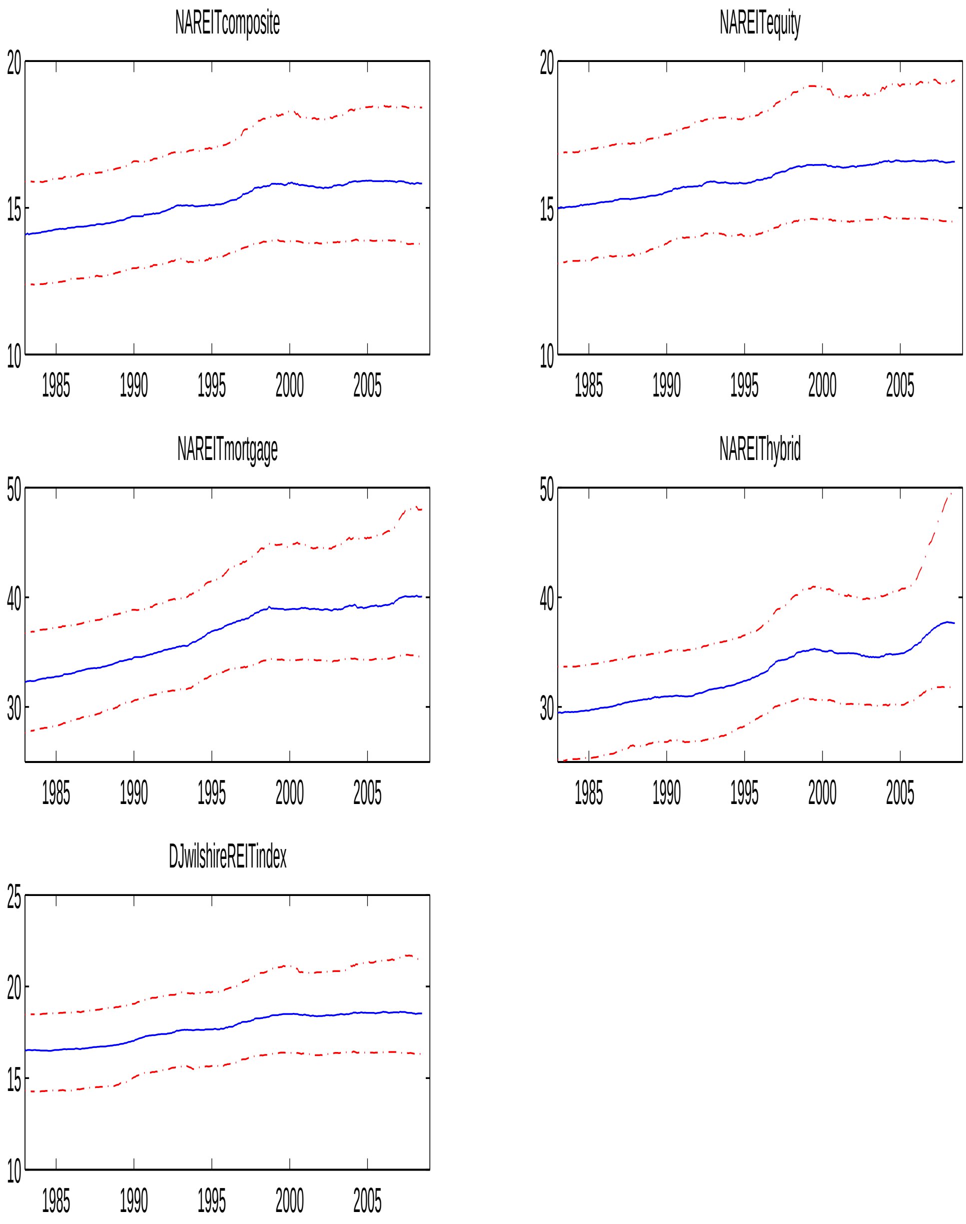

Figure 4: Sensitivity to the intercept – Posterior Medians plus the 90% Bayesian confidence intervals.

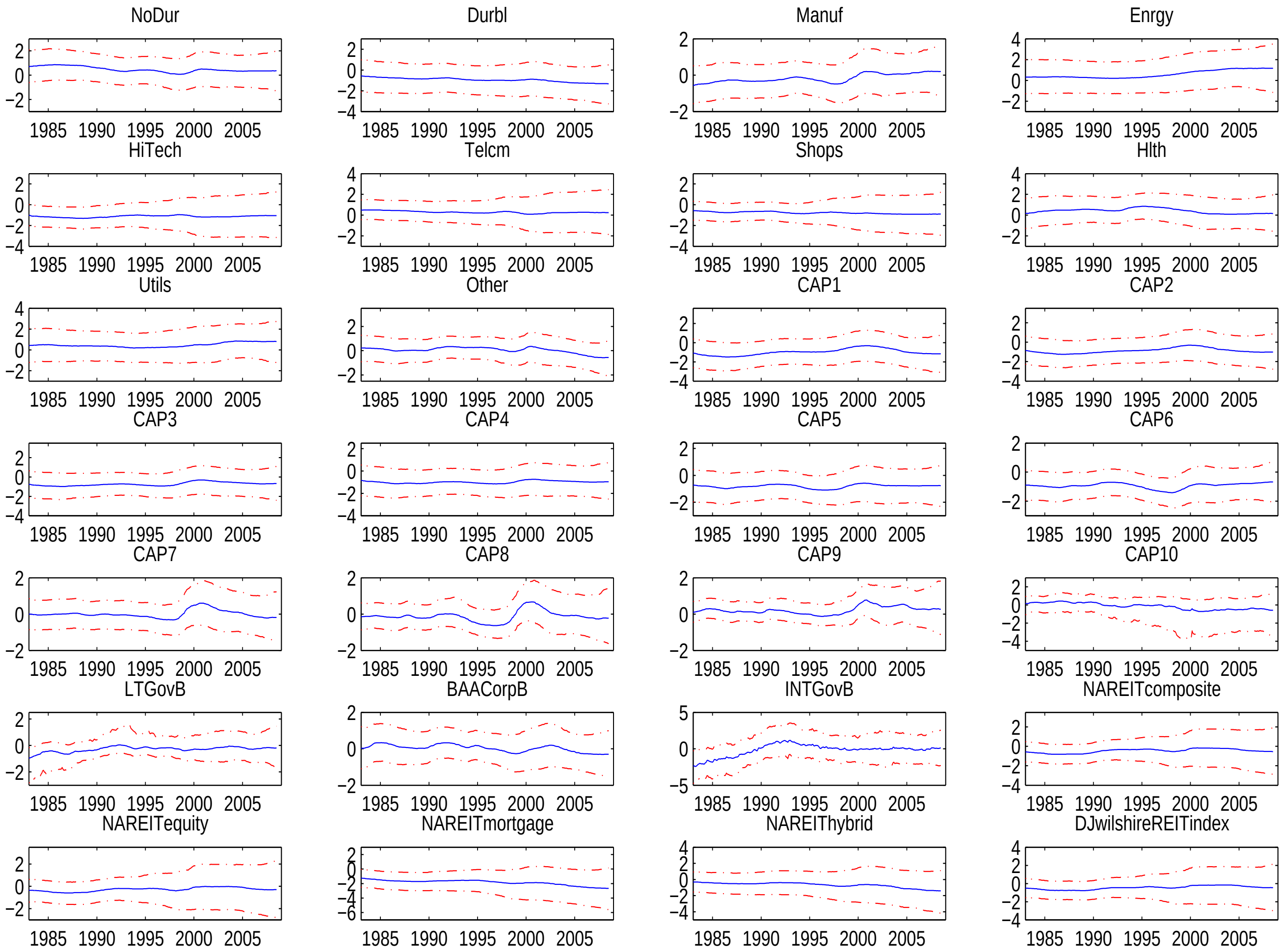

Figure 5: Sensitivity to VW – Posterior Medians plus the 90% Bayesian confidence intervals.

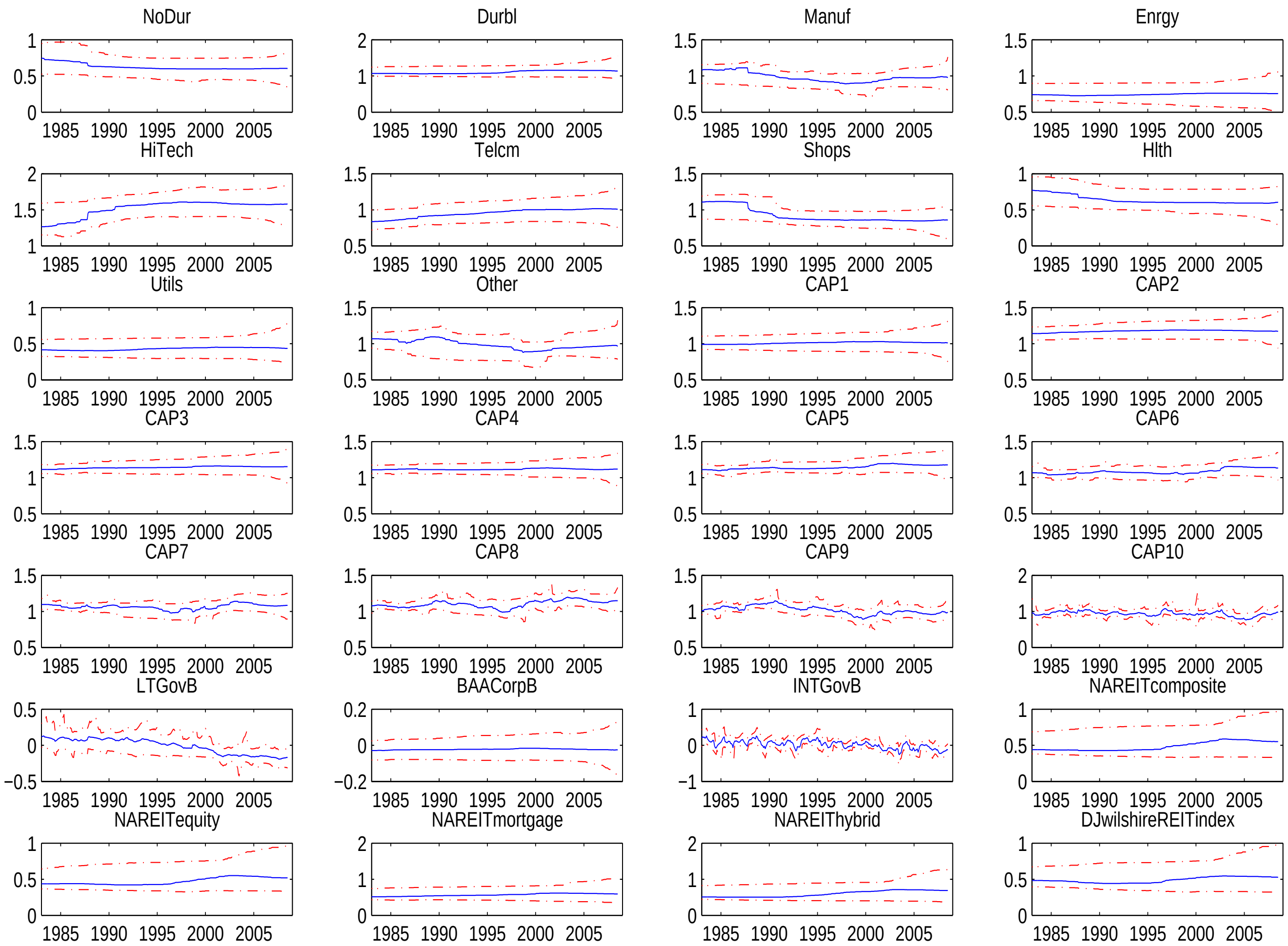

Figure 6: Sensitivity to PREM – Posterior Medians plus the 90% Bayesian confidence intervals.

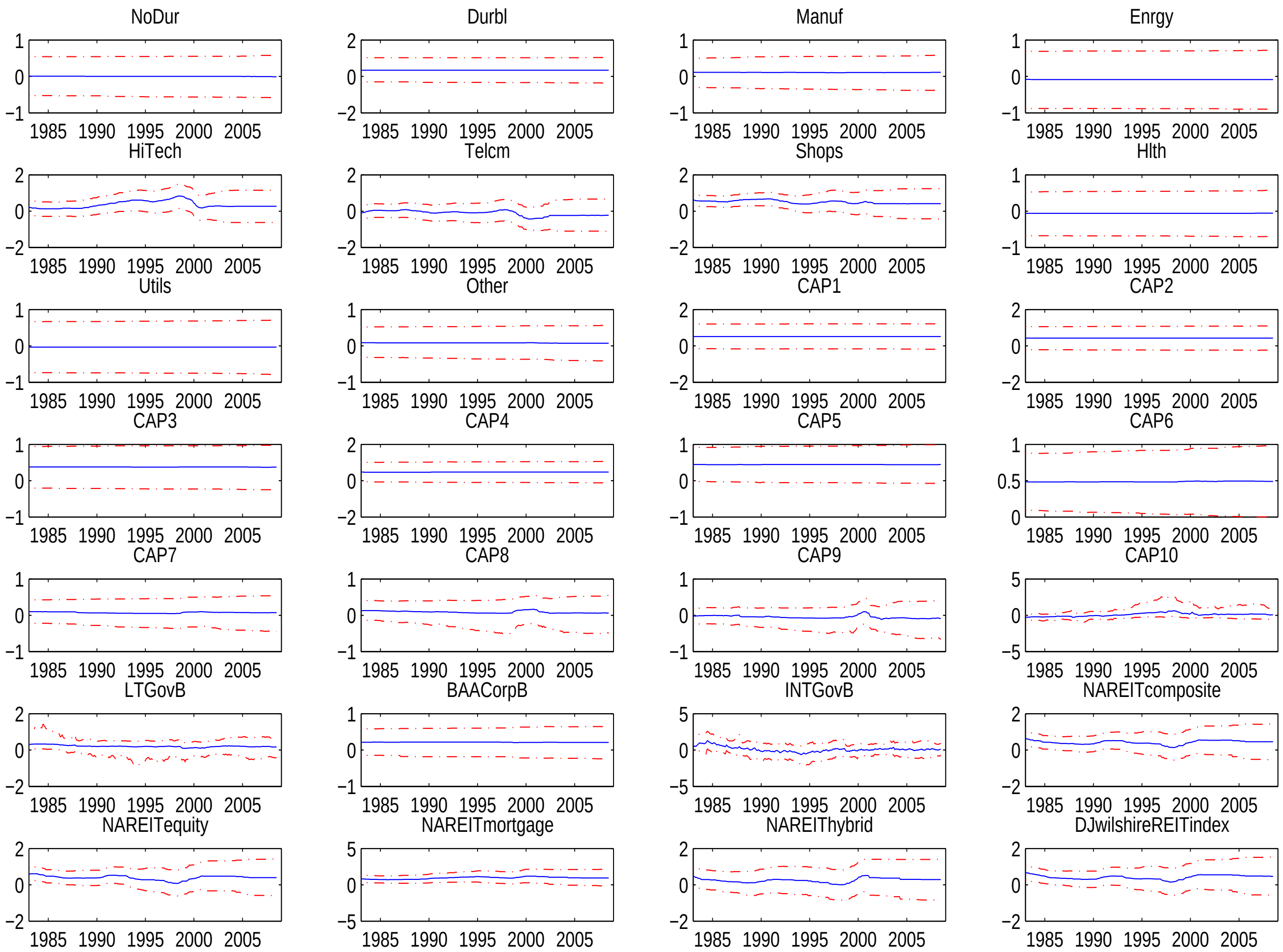

Figure 7: Sensitivity to DSLOPE – Posterior Medians plus the 90% Bayesian confidence intervals.

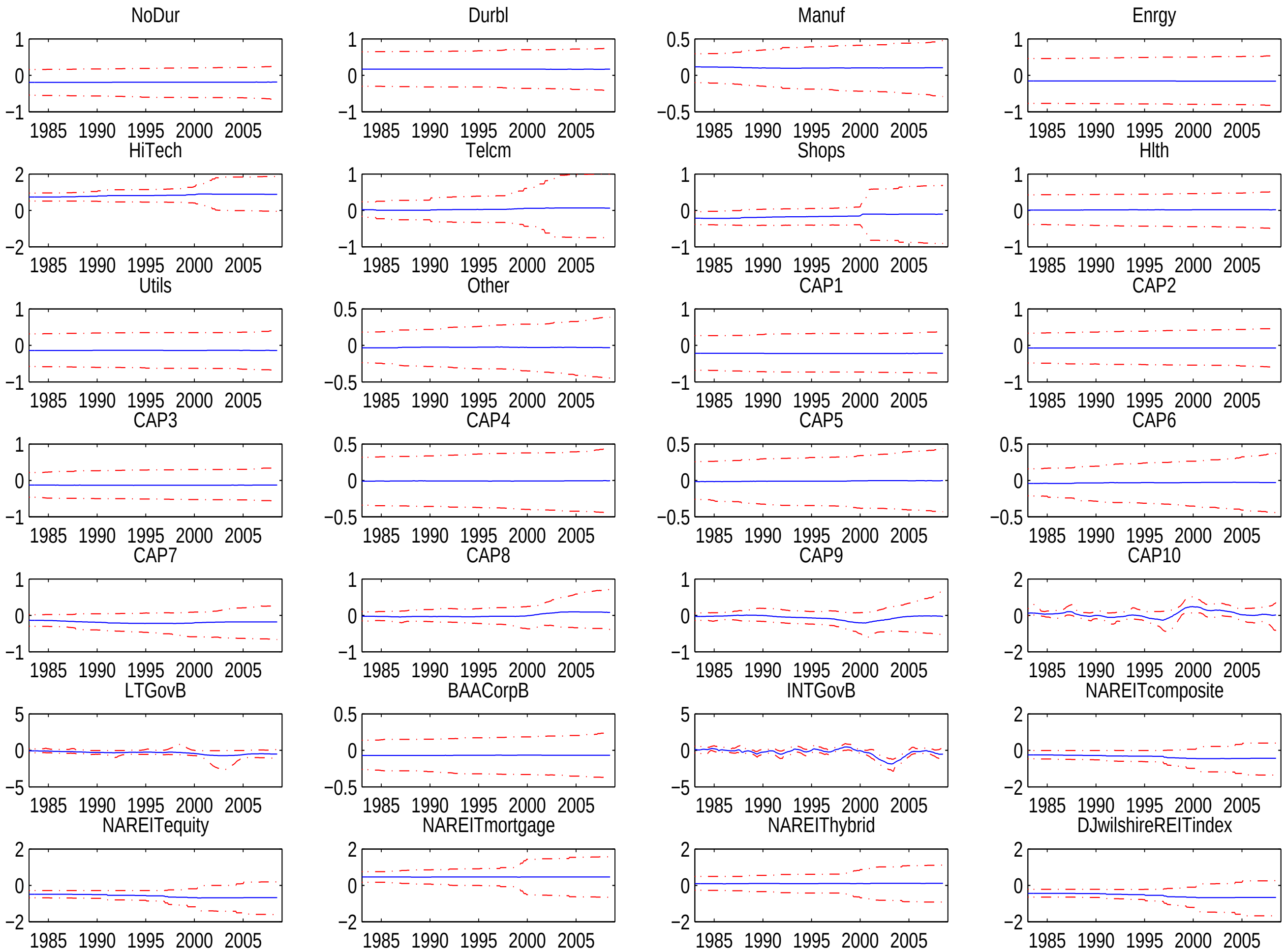

Figure 8: Sensitivity to UI – Posterior Medians plus the 90% Bayesian confidence intervals.

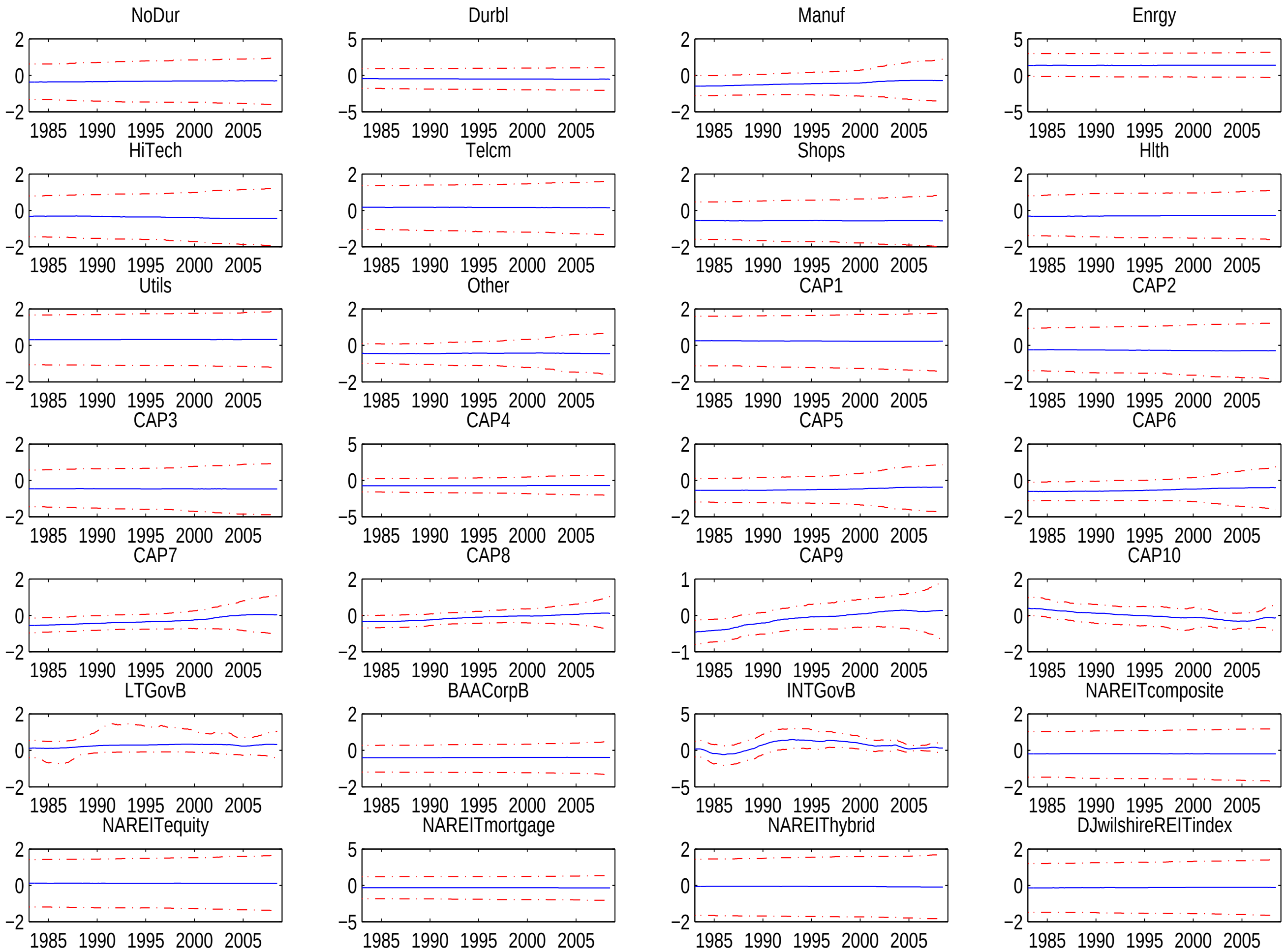

Figure 9: Sensitivity to IP – Posterior Medians plus the 90% Bayesian confidence intervals.

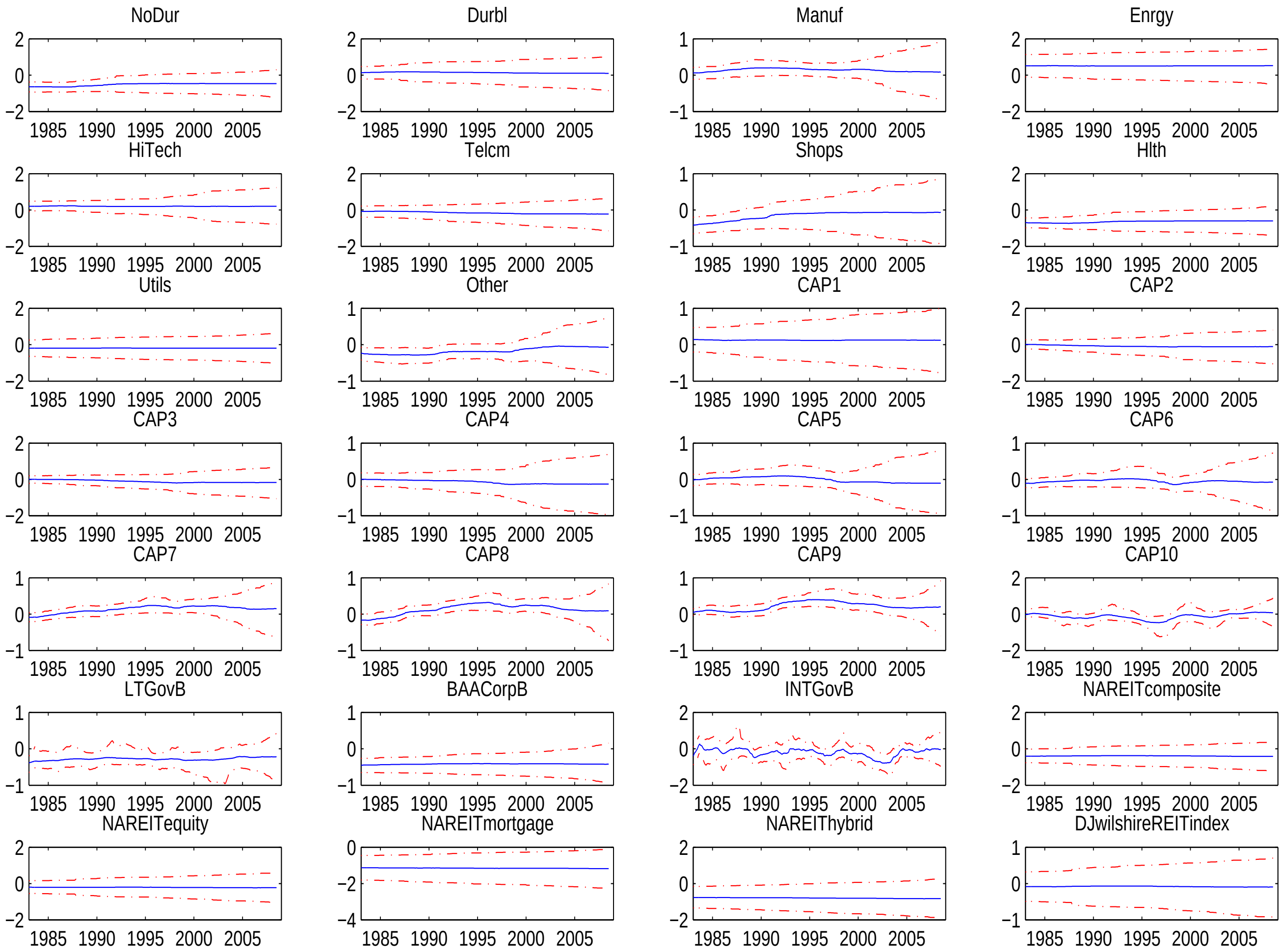

Figure 10: Sensitivity to CONSUM – Posterior Medians plus the 90% Bayesian confidence intervals.

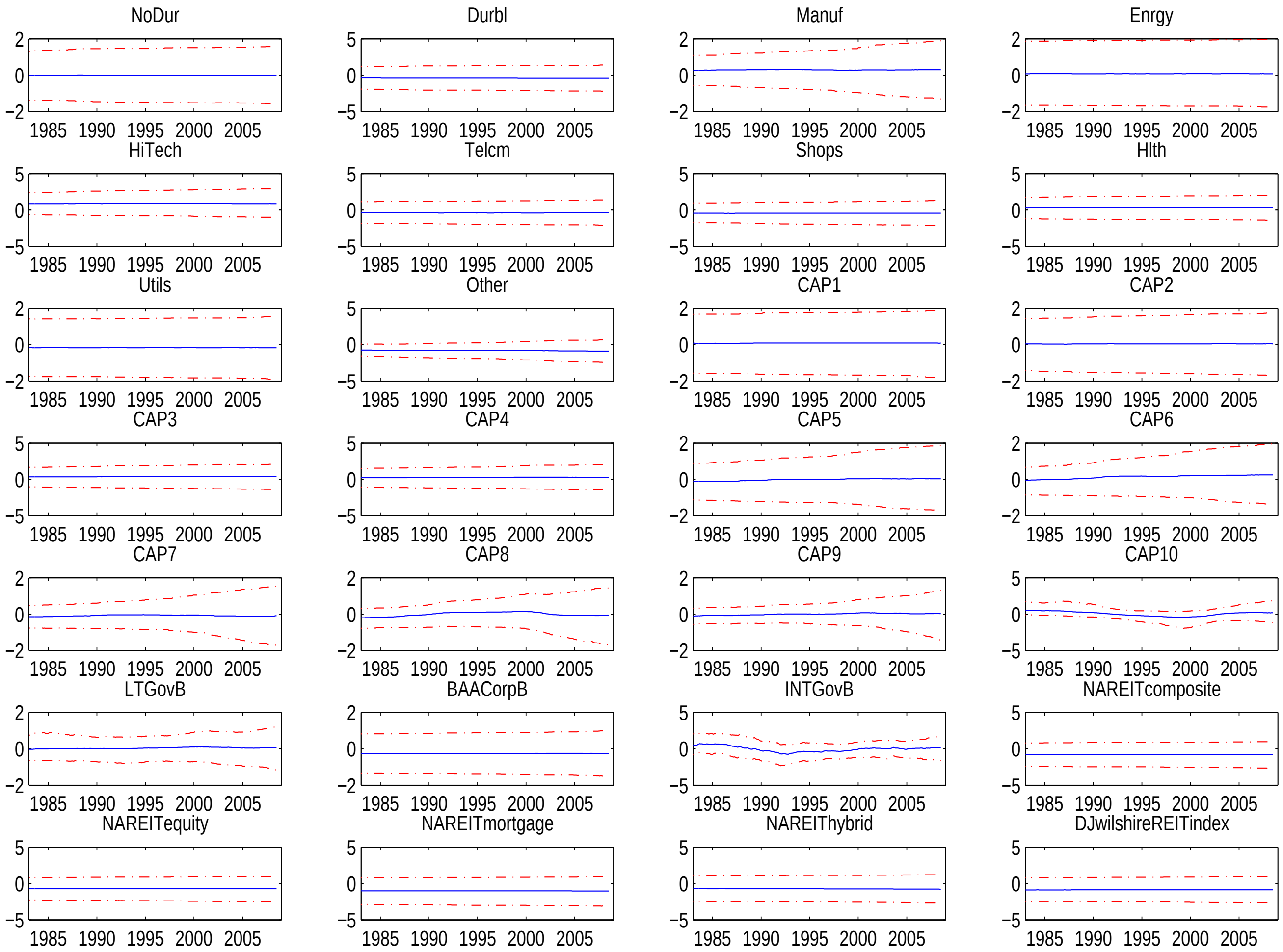

Figure 11: Sensitivity to REALTB – Posterior Medians plus the 90% Bayesian confidence intervals.

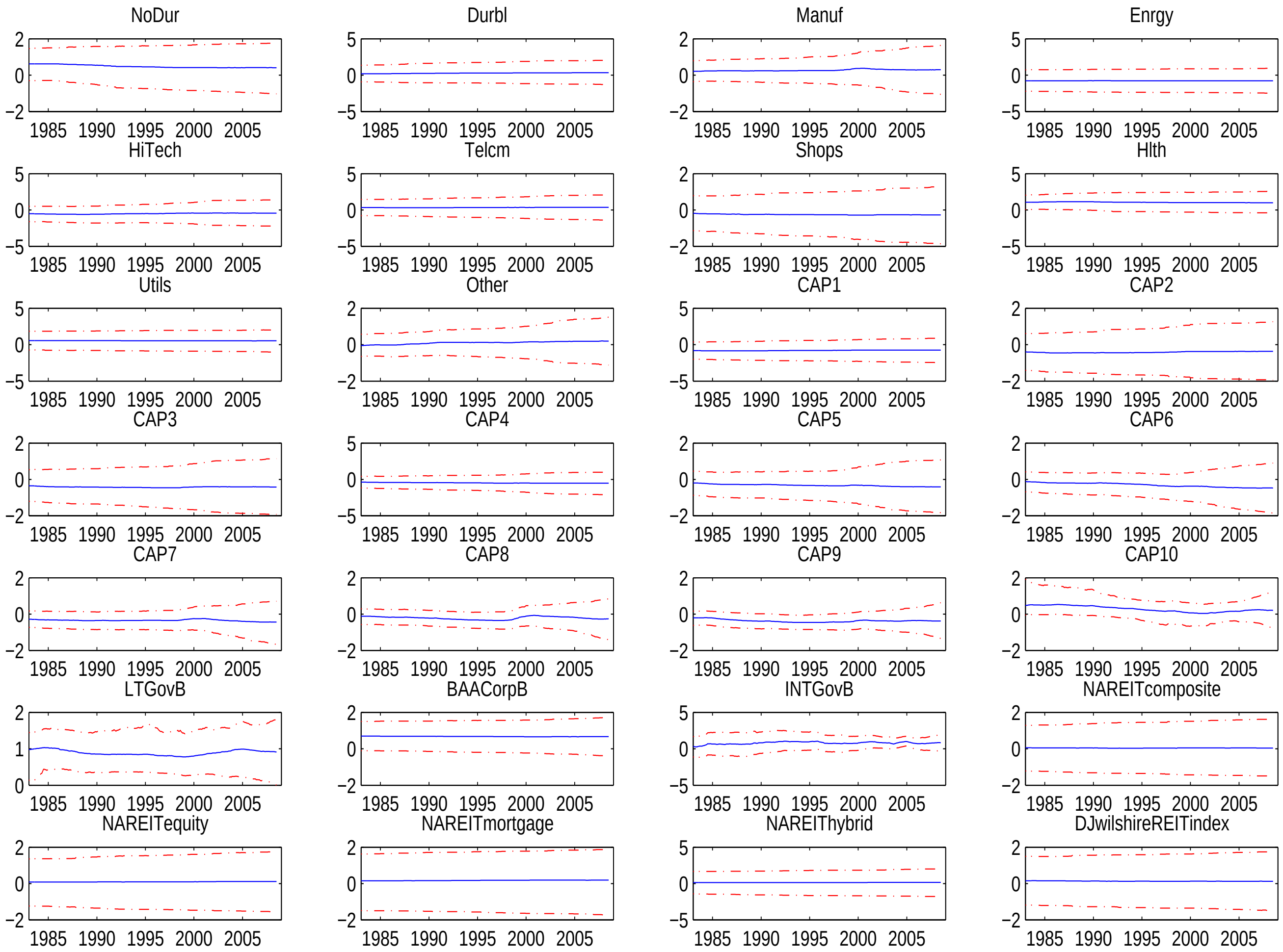

Table 3: Average Risk Premiums *à* la Ouysse and Kohn (2009)

Period	*ZBeta*	*VW*	*Prem*	*DSlope*	*UI*	*IP*	*Consum*	*RealTb*
1983:01-2008:07	**0.374**	**0.232**	**0.02**	**-0.15**	**-0.01**	**0.06**	**0.06**	**-0.01**
5th Perc.	−2.88	−7.14	−2.67	−3.73	−1.28	−2.79	−0.74	−1.24
95th Perc.	4.02	7.15	2.89	3.86	1.17	3.55	0.86	1.11
Std. error	0.12	0.26	0.11	0.14	0.044	0.11	0.03	0.05
Num. Std. error	0.15	0.23	0.14	0.14	0.03	0.18	0.04	0.06
1983:01-1992:12	**0.26**	**0.39**	**-0.16**	**-0.51**	**-0.07**	**-0.17**	**-0.014**	**0.09**
5th Perc.	−2.78	−6.71	−1.735	−4.07	−0.78	−2.35	−0.64	−0.70
95th Perc.	3.72	8.17	1.61	2.95	0.65	2.245	0.58	0.87
Std. error	0.18	0.46	0.10	0.2	0.043	0.13	0.03	0.04
Num. Std. error	0.2	0.27	0.12	0.23	0.04	0.04	0.048	0.04
1998:01-2008:07	**0.42**	**-0.08**	**0.15**	**0.01**	**0.04**	**0.34**	**0.13**	**-0.09**
5th Perc.	−3.03	−8.43	−3.88	−3.72	−1.51	−3.08	−0.92	−1.58
95th Perc.	4.09	7.15	4.36	3.73	1.69	3.93	1.06	1.82
Std. error	0.2	0.42	0.23	0.22	0.09	0.21	0.05	0.11
Num. Std. error	0.3	0.43	0.3	0.19	0.054	0.37	0.07	0.11

Notes: The definition of the economic risk variables is in Table 1. The model being estimated is the linear regression in (4). We report average values together with the 5th and the 95th percentiles of the median risk premiums for the entire sample and two subperiods, respectively. The standard errors for the average of the risk premiums over time under the assumption of no autocorrelation and the numerical standard errors are also reported.

Table 4: Average Risk Premiums - Bayesian version of the Fama-MacBeth (1973) statistic

Period	$ZBeta$	VW	$Prem$	$DSlope$	UI	IP	$Consum$	$RealTb$
1983:01-2008:07	**0.376**	**0.232**	**0.013**	**-0.155**	**-0.007**	**0.0606**	**0.058**	**-0.005**
2.5th Perc.	0.13	−0.02	−0.22	−0.5	−0.16	−0.21	−0.08	−0.19
5th Perc.	0.17	0.02	−0.183	−0.452	−0.133	−0.172	−0.067	−0.154
95th Perc.	0.579	0.436	0.224	0.162	0.116	0.262	0.182	0.126
97.5th Perc.	0.62	0.48	0.27	0.24	0.14	0.30	0.21	0.15
1983:01-1992:12	**0.258**	**0.398**	**-0.162**	**-0.514**	**-0.07**	**-0.16**	**-0.017**	**0.091**
2.5th Perc.	−0.17	0.03	−0.53	−1.10	−0.32	−0.59	−0.24	−0.19
5th Perc.	−0.09	0.09	−0.48	−1.003	−0.29	−0.53	−0.20	−0.16
95th Perc.	0.59	0.7	0.16	0.007	0.134	0.16	0.175	0.33
97.5th Perc.	0.67	0.78	0.24	0.11	0.18	0.23	0.21	0.38
1998:01-2008:07	**0.42**	**-0.075**	**0.145**	**-0.002**	**0.043**	**0.336**	**0.127**	**-0.083**
2.5th Perc.	0.02	−0.55	−0.3	−0.57	−0.21	−0.12	−0.12	−0.34
5th Perc.	0.082	−0.47	−0.228	−0.47	−0.17	−0.037	−0.086	−0.302
95th Perc.	0.786	0.285	0.54	0.539	0.238	0.71	0.354	0.12
97.5th Perc.	0.86	0.35	0.62	0.65	0.27	0.77	0.41	0.17

Notes: The definition of the economic risk variables is in Table 1. The model being estimated is the linear regression in (4). We construct a time series of estimated risk premiums for each iteration i and compute accordingly the average risk premiums: as a result, a posterior distribution is easily obtained. We report its median values together with the 2.5th, 5th, 95th and the 97.5th percentiles.

Table 5: Average Risk Premiums - Median Betas

Period	*ZBeta*	*VW*	*Prem*	*DSlope*	*UI*	*IP*	*Consum*	*RealTb*
1983:01-2008:07	**0.43**	**0.18**	**0.01**	**-0.29**	**-0.03**	**0.14**	**0.15**	**0.14**
5th Perc.	-3.55	-7.42	-5.44	-5.61	-2.51	-3.95	-2.26	-2.93
95th Perc.	4.96	7.48	6.16	5.37	2.28	4.97	2.95	3.42
Std. error	0.15	0.29	0.20	0.19	0.085	0.15	0.1	0.12
1983:01-1992:12	**0.3**	**0.40**	**-0.52**	**-0.92**	**-0.10**	**-0.13**	**-0.02**	**0.38**
5th Perc.	-3.91	-7.47	-5.16	-6.59	-2.54	-3.91	-2.12	-3.83
95th Perc.	5.37	9.14	3.59	4.12	2.41	3.5	2.39	3.71
Std. error	0.26	0.51	0.25	0.3	0.13	0.20	0.12	0.21
1998:01-2008:07	**0.48**	**-0.17**	**0.36**	**0.03**	**-0.01**	**0.5**	**0.29**	**-0.015**
5th Perc.	-3.5	-9.246	-6.08	-4.72	-3.29	-4.13	-2.83	-2.35
95th Perc.	4.88	7.48	7.83	5.47	2.56	6.0	3.93	3.28
Std. error	0.22	0.45	0.39	0.30	0.14	0.27	0.18	0.19

Notes: The definition of the economic risk variables is in Table 1. The model being estimated is the linear regression in (4). We report average values together with the 5th and the 95th percentiles of the median risk premiums for the entire sample and two subperiods, respectively, obtained by using the median betas from the fist pass as regressors. The standard errors for the average of the risk premiums over time are also reported.

Table 6: Decomposition of the predictable variation of monthly portfolio returns (1983:01 - 2008:07) by economic risk variables.

	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

Notes: The first two ratios are defined in (5). VR1 denotes the ratio of the variance of the multifactor model's predicted returns from the economic premiums and the asset's betas relative to the variance of the expected returns from a linear regression on the instrumental variables. VR2 denotes instead the ratio of the variance of the predictable part of a return that is not explained by the multifactor model relative to the variance of the expected returns. The other ratios are computed as explained at the end of section 3. The variable "Int Eff" reflects the fact that the total predictable variation in the set of economic risk variables is different than sum of the predictable variation due to each of the variables individually.

Table 7: Average Risk Premiums *à* la Ouysse and Kohn (2009) without REITs

Period	*ZBeta*	*VW*	*Prem*	*DSlope*	*UI*	*IP*	*Consum*	*RealTb*
1983:01-2008:07	**0.33**	**0.25**	**0.07**	**-0.1**	**-0.036**	**0.01**	**-0.01**	**0.016**
5th Perc.	−2.6	−4.52	−1.84	−5.88	−1.56	−3.32	−0.82	−1.59
95th Perc.	3.4	4.05	1.96	5.25	1.42	3.66	0.88	1.74
Std. error	0.11	0.16	0.071	0.20	0.05	0.13	0.03	0.064
Num. Std. error	0.11	0.12	0.035	0.20	0.048	0.15	0.04	0.06
1983:01-1992:12	**0.23**	**0.37**	**0.01**	**0.14**	**-0.04**	**-0.25**	**-0.05**	**0.04**
5th Perc.	−2.55	−4.51	−1.51	−6.31	−0.89	−2.49	−0.79	−0.816
95th Perc.	3.24	4.73	1.65	5.24	0.82	1.97	0.61	0.81
Std. error	0.16875	0.28398	0.10487	0.34861	0.054761	0.12379	0.036246	0.048185
Num. Std. error	0.18	0.17	0.06	0.36	0.04	0.13	0.023	0.03
1998:01-2008:07	**0.21871**	**0.14934**	**0.1184**	**-0.43416**	**-0.085451**	**0.076009**	**0.028043**	**0.038609**
5th Perc.	−3.03	−4.715	−2.80	−5.31	−1.72	−5.12	−0.86	−2.65
95th Perc.	3.78	3.98	2.72	3.68	2.07	5.2	1.08	2.60
Std. error	0.18	0.24	0.13	0.28	0.11	0.26	0.06	0.14
Num. Std. error	0.14	0.26	0.036	0.204	0.1	0.27	0.09	0.15

Notes: The definition of the economic risk variables is in Table 1. The model being estimated is the linear regression in (4). We report average values together with the 5th and the 95th percentiles of the median risk premiums for the entire sample and two subperiods, respectively. The standard errors for the average of the risk premiums over time under the assumption of no autocorrelation and the numerical standard errors are also reported.

Table 8: Average Risk Premiums - Bayesian version of the Fama-MacBeth (1973) statistic without REITs

Period	*ZBeta*	*VW*	*Prem*	*DSlope*	*UI*	*IP*	*Consum*	*RealTb*
1983:01-2008:07	**0.33**	**0.26**	**0.07**	**-0.097**	**-0.04**	**0.01**	**-0.01**	**0.017**
2.5th Perc.	0.036	−0.025	−0.23	−0.57	−0.22	−0.32	−0.18	−0.16
5th Perc.	0.082	0.024	−0.18	−0.49	−0.18	−0.27	−0.15	−0.135
95th Perc.	0.56	0.49	0.31	0.3	0.09	0.284	0.125	0.18
97.5th Perc.	0.61	0.54	0.37	0.11	0.37	0.33	0.15	0.21
1983:01-1992:12	**0.23**	**0.37**	**0.01**	**0.16**	**-0.04**	**-0.26**	**-0.05**	**0.04**
2.5th Perc.	−0.22	−0.027	−0.404	−0.624	−0.31	−0.70	−0.29	−0.25
5th Perc.	−0.17	0.044	−0.316	−0.5	−0.26	−0.63	−0.246	−0.21
95th Perc.	0.59	0.72	0.35	0.76	0.16	0.12	0.13	0.325
97.5th Perc.	0.66	0.79	0.43	0.9	0.19	0.2	0.17	0.38
1998:01-2008:07	**0.21**	**0.16**	**0.117**	**-0.44**	**-0.085**	**0.08**	**0.028**	**0.038**
2.5th Perc.	−0.28	−0.39	−0.48	−1.22	−0.4	−0.597	−0.29	−0.3
5th Perc.	−0.2	−0.29	−0.38	−1.08	−0.34	−0.47	−0.23	−0.23
95th Perc.	0.635	0.6	0.62	0.21	0.15	0.63	0.29	0.32
97.5th Perc.	0.703	0.69	0.72	0.33	0.18	0.73	0.36	0.39

Notes: The definition of the economic risk variables is in Table 1. The model being estimated is the linear regression in (4). We construct a time series of estimated risk premiums for each iteration i and compute accordingly the average risk premiums: as a result, a posterior distribution is easily obtained. We report its median values together with the 2.5th, 5th, 95th and the 97.5th percentiles.

A Bayesian investigation of the predictive power of the yield spread for economic activity

Abstract

The term structure of interest rates is considered to vary with the business cycle and, hence, to contain useful information to predict measures of economic activity. In particular, many empirical studies look at the yield spread, the difference between the yields on long-term and short-term Treasury securities, and find that it helps forecast output growth especially at short-medium horizons. Yet, this predictive power is documented to have declined since the mid-eighties. The present work aims to analyze its time varying nature for the United States in light of the recent economic events. To this end, I follow Benati and Goodhart (2008) and derive a measure of (*in-sample*) marginal predictive power at horizons of four and eight quarters from bivariate and fourvariate time-varying VARs. The evidence confirms the findings in the previous literature and also suggests that such predictive power has not changed much lately, apart from a short-lived hike around the time of the recession in 2001. Furthermore, it appears to be more important the contribution of inflation and the short rate.

1 Introduction

The yield curve is considered to be a reliable predictor of both inflation and output over time. Notwithstanding, a leading theoretical framework is still missing. Indeed, Wheelock and Wohar (2009) start a recent survey of the literature from a telling quotation of Benati and Goodhart (2008) about the predictive power of the yield spread: a "stylized fact in search of a theory".[1] Absent a dominant model to refer to, it comes very natural to snoop the entire yield curve in search for a good predictor.

In particular the term spread, measured as the difference between the yields on long-term and short-term Treasury securities, has always received much attention. It is widely recognized that its predictive content for output growth has diminished in the nineties: yet, none of the studies in the literature has explictly modelled the time variation in the forecasting equation parameters. The only exception is Benati and Goodhart (2008) (henceforth BG) who estimate, among other things, a time varying parameter VAR with stochastic volatility for output growth, inflation, short rate and term spread. Their aim is to assess the *marginal* contribution of the yield spread and discriminate between the two main candidate explanations (the monetary policy-based one and Harvey's real yield curve's) for such predictive power.

Needless to say, all the literature predates the current economic crisis. Therefore, I reconsider the main aspects in light of the more recent events. For instance, the yield curve inversion in August 2006 was correctly foretelling an upcoming recession but, at that time, most reinforced their beliefs about the actual forecast breakdown. So an interesting question to ask is whether the signal in the yield spread is back. To this end, I use a very similar approach to BG, i.e. a time-varying parameters VAR with stochastic volatility *and* correlations.[2]

[1] As the authors remark, much of the empirical literature has not tried to discriminate between the main competing explanations. Empirical and theoretical approaches essentially sway in that the former analyze the nominal yield curve while the latter the real one. Other differences involve the fact that policymakers are tipically postulated to care about measures of output gap rather than its plain level. For a brief review please see the next section.

[2] For a lightening and thoughtful analysis on how to model instability see Cogley (2005). It is worth noting that reasonable alternative models may opt for a finite-state Markov representation (more generally, hidden Markov models) whereas here parameters and conditional second moments are forced to move continuously. Furthermore, in most applications with regime switching evidence is found in favour of more than two regimes which on the other hand lack straight economic interpreation.

Likewise, it is important to understand the determinants of its behaviour over time. Firstly, I analyze the robustness of the results when other variables are also taken into account. For instance, in line with some previous papers (notably BG and An, Piazzesi and Wei (2006)) I include the short rate on both theoretical and empirical grounds. In a nutshell, I compute a measure of marginal contribution of the yield spread based on the difference between the fit for output growth of a fourvariate VAR and a trivariate VAR (in other terms, a restricted fourvariate VAR), both specifications including inflation as well. The scope of the search can be further extended by analyzing yield factors other than slope like level and curvature, whether estimated or retrieved as in Mumtaz and Surico (2009). The new candidate variables or measures can be evaluated on a one-by-one basis or, more elegantly, in a framework that endogenously account for model uncertainty and variable selection.[3]

BG notice that the marginal predictive power markedly increases in periods of higher uncertainty about the monetary policy regime, when risk premia plausibly play some role. De Pace (2009) directly explores the link between the fall in the leading properties of the term spread and the variation in inflation risk. The author estimates the marginal processes of both long term yield bonds and the yield spread on the ground that the former's volatility proxies for inflation risk while the latter incorporates information on inflation expectations: the evidence shows the importance of breaks that have brought about a decrease in conditional variances.In this respect, I intend to explore the role of uncertainty by including the default premium as measured by the BAA-AAA spread. For example, a tentative analysis could be based on the correlation exhibited by the default premium and the measure of marginal contribution.

To summarize, the results show that the predictive power of the yield spread at the horizon of 4 quarters has peaked in the mid-eighties and has since decreased as reported in many studies and that, apart from a very short-lived episode related to the recession of 2001, it has remained low up to the second quarter of 2009. The 8-quarter ahead predictive power, instead, is very low and shows no sign of significant variation.

[3]An example of the latter approach is Koop, Jochmann and Strachan (*forthcoming*) who in turn apply the stochastic search variable selection method, an automatic model selectiond device, to reduce over-parameterization problems.

The remainder of the paper is organized as follows. Section 2 briefly resumes the literature, Section 3 describes the data and Section 4 presents the model structure, discusses the prior choices and illustrates the algorithm used for posterior simulation. Section 5 presents the results while Section 6 concludes.

2 Literature Review

Why does the yield spread contain information about future output growth? Surprisingly, few papers have addressed this question. Estrella (2004) [see also Estrella and Trubin (2006)] points out that the relationship with output and inflation is not structural but reflect monetary policy decisions, in other words the reaction function. There are two polar scenarios: if the monetary authority reacts only to output fluctuations and focuses just on changes in the interest rate, the yield curve is the optimal predictor whereas such predictive power disappears in case reactions to both inflation and output approaches infinity. In between the two cases, information in the yield curve can be combined with other data.

Alternatively, according to an explanation firstly stated by Harvey (1988), it is the *real* term structure that is linked to measures of future economic activity through intertemporal consumption smoothing. Bordo and Haubrich (2004, 2008) find that inflation persistence is key to forecasting success because of the role played by expectations. Low persistence acts as a noise in the *nominal* term spread on the ground that real shocks usually do not equally impact on both ends of the yield curve.

Several authors recognize a remarkable reduction in the forecast uncertainty shown by naive models in the last period of macroeconomic stability whereas D'Agostino, Giannone and Surico (2007) add that there is also a significant decline in the forecasting gain of sophisticated models, in particular the ones exploiting large information datasets.[4] The main reason is the presence of

[4]Giannone, Lenza and Reichlin (2008) offer many insights about the causes of the Great Moderation. They contend the majority view supporting the "good luck" hypothesis for output. In a large scale model, instead, are changes in propagation that seem to be the main mechanism at work: for the problem at hand, correlations between GDP and other variables have dropped causing a decline in both the sample variability and predictability. Similar is the message in Benati and Surico (2009) who warns that in typical structural VARs' exercises "good policy" may be easily mistaken for "good luck".

changes in dynamic correlations between variables rather than in the autocorrelations of output and inflation: apart from the Phillips curve case, the authors document a breakdown in the relationship between the yield-curve and output growth.[5]

Estrella, Rodrigues and Schich (2003) analize the stability of the predictive content of the yield curve for Germany and United States: for U.S. they find weak evidence of a break around September 1983 with a one-year horizon while the evidence disappears as the horizon increases. Giacomini and Rossi (2006) build on Estrella *et al.* (2003) for output growth forecasting and favour of the concept of forecast breakdown. It encompasses the issue of coefficient stability along with, say, forecast error distributions: the focus becomes now the reliability of a given model for forecasting or, in other terms, the comparison of fitted errors and the out-of-sample forecast errors according to a quadratic loss function. Forecast breakdown appears to be robust to the different tests employed and strictly related to monetary policy changes. In particular, it increases during both the Burns-Miller and Volker periods whereas the Greenspan era has seen this relationship stabilizing. The authors report an explanation that relies on the role of the decreased volatility of output growth in making in-sample and out-of-sample losses closer. Results are anyway not at odds with the dominant literature attributing the great bulk of predictability to the pre-1984 period as the stabilization during the Greenspan era, following the high surprise losses in the first Volker period, implies a reduced forecasting ability with respect to the early sample.[6]

Some works have recently decomposed this quantity in its two theoretical constituents, namely the expectation component and the term premium component, on the ground that it is the former one that very likely foretells future movements in economic activity. The first paper in this context is Hamilton and Kim (2002). Then Favero, Kaminska and Soderstrom (2005) who, on the contrary, use a real-time VAR to compute short term rate expectations. In both cases a lower term premium turns out to predict slower GDP growth.

[5]A clear message is that inference based on a long sample is basically influenced by the overperfomance before 1985. D'Agostino et al. (2007) also confirm Stock and Watson (2003) in that univariate models of real activity show little changes over this two periods.

[6]The authors explicitly assert that "detection of a forecast breakdown does not necessariily mean that the model should be discarded. Rather, it means that the model's future performance will likely not be consistent with its past performance".

At times in which the term premium is small and the yield curve is relatively flat, small changes in expectations may reduce the accuracy of recession forecasts. This is borne out by the evidence in Rosenberg and Maurer (2008) who apply the Kim and Wright (2005) measure and comment that the term premium follows a longer trend with a decline since 1984 while the expectation component mimics very well the dynamics of the term spread. As for in-sample fit, the authors report a better performance for the model featuring the two components separately even though the term premium does not result significant.[7] The out-of-sample results do not differ much. Rudebusch, Sack and Swanson (2007) instead find that the coefficient on the term premium is negative and marginally statistically significant: a decline in the term premium is associated with future growth. These sorts of exercises are, however, limited by the fact that term premia cannot be estimated very precisely and theoretical models do not provide a clear guidance as to the relationship between yields and economic activity. The main difficulty is how to correctly extract the two unobservable components: many alternatives are available but none of them appears superior to the others.[8] As a matter of fact, Kim and Orphanides (2007) use survey forecasts as a proxy for market's expectations of future short term rates.

Finally, Ang Piazzesi and Wei (2006) challenge the conventional practice of using the yield-spread for GDP forecasts and show that the short rate is the dominant source of information in a term structure with no-arbitrage restrictions.[9] Typical unrestricted OLS regressions instead tend to erroneously emphasize the role of the term spread as they do not efficiently capture the information in the cross-section of yields. On the other hand, Zhu and Rahman (2009) estimate a regime switching macro-finance model based on a dynamic Nelson-Siegel setup without no-arbitrage restrictions and the impulse responses suggest that future levels of capacity utilization are more sensitive to slope factor shocks rather than level factor shocks, which are found to be highly correlated with long yields.

[7] In other terms, when using the yield spread alone one is essentially imposing the restriction of equality in coefficients. Assuming the decomposition is carried out not too imprecisely, the evidence suggests that the restriction does not hold.

[8] A brilliant coverage is available in Rudebusch, Sack and Swanson (2007).

[9] The short rate is found to be highly correlated with the Expectations Hypothesis (EH) spread, in a similar vein to Rosenberg and Maurer (2008).

3 Data

The data are quarterly and cover the period from 1953 to the second quarter of 2009 but the first ten years of observations are used for prior calibration. In short, I consider four variables: real GDP growth (percentage variation over the previous quarter), CPI inflation rate (annualized variation with respect to the last month of the previous quarter), the short term rate (the 3-month Treasury bill rate, secondary market) and the yield spread computed as the 10-year government bonds' yield minus the 3-month Treasury bill rate. I finally consider the default spread, measured as the yield spread between BAA and AAA rated corporate bonds. The source is FRED, the database maintained by the Federal Reserve Bank of St. Louis.

The four series are plotted together in Figure 1 for the actual sample under consideration in the next section, that is from 1963:I to 2009:II. There is visual evidence of changing characteristics in the processes governing the evolution of these variables. Firstly it is apparent the decrease in variability shown by both output growth and inflation in the second half of the sample, this phenomenon being referred to in the literature as the "Great Moderation". The short term interst rate has inverted its upward trend since then, while the yield spread shows an increased range of variability from the late seventies. Finally, note the tremendous effect the onset of the financial crisis has exerted on output growth and inflation, notably, and on interest rates, to a lesser extent. The use of a multivariate model with time-varying parameters seems then appropriate to capture these features of the data.

4 A TVP-VAR with Stochastic Volatility

4.1 The empirical model

In line with the current prevailing literature in macroeconometrics, a time-varying parameters VAR(p) model is used:

$$Y_t = B_{0,t} + B_{1,t}Y_{t-1} + ... + B_{p,t}Y_{t-p} + u_t \equiv X_t^{\prime}\theta_t + u_t \tag{1}$$

where Y_t is a vector that collects the observed endogenous variables.[10] For consistency with similar works and, more importantly, for computational feasibility the lag order is set to two ($p = 2$).[11] As for the VAR's reduced form innovations in (1), $u_t \sim N(0, \Omega_t)$ where Ω_t is factored as

$$\Omega_t = A_t^{-1}\Sigma_t\Sigma_t^{'}(A_t^{-1})^{'} \tag{2}$$

with Σ_t and A_t defined as[12]

$$\Sigma_t \equiv \begin{bmatrix} \sigma_{1,t} & 0 & \cdots & 0 \\ 0 & \sigma_{2,t} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_{N,t} \end{bmatrix}, A_t \equiv \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \alpha_{21,t} & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{N1,t} & \alpha_{N2,t} & \cdots & 1 \end{bmatrix}. \tag{3}$$

It follows that

$$Y_t = X_t^{'}\theta_t + A_t^{-1}\Sigma_t\epsilon_t \tag{4}$$

where $\epsilon_t \sim N(0, I_n)$.

The time-varying parameters, collected in the vector θ_t, evolve in the following way:

$$p(\theta_t|\theta_{t-1}, Q) = I(\theta_t)f(\theta_t|\theta_{t-1}, Q) \tag{5}$$

with $I(\theta_t)$ being an indicator function acting as a reflecting barrier in case of unstable draws.[13] On the other hand, $f(\theta_t|\theta_{t-1}, Q)$ is characterized by the following law of motion

[10] The setup is here described in general terms whereas in the following a range of several specifications, from univariate to fourvariate autoregressive models, is estimated.

[11] Most, if not all, of the modelling assumptions in this part are made according to common practice in the literature. In the sequel, only deviations from it will be remarked and justified.

[12] The structure of A_t is also an identification scheme. Whenever the short term rate is involved, it is ordered as last in the system so as to single out the monetary policy shock. However, this aspect is not of primary interest for the present analysis.

[13] The stability constraint is imposed for both theoretical and practical reasons. As for the latter, by this way it is possible to apply frequency domain techniques at each point in time and derive measures of persistence and predictability. Anyway, for more reflections on the use of such a constraint see section 5.

$$\theta_t = \theta_{t-1} + \eta_t \tag{6}$$

with $\eta_t \sim N(0, Q)$. In turn, the $\sigma_{i,t}$'s evolve according to geometric random walks

$$ln\sigma_{i,t} = ln\sigma_{i,t-1} + \nu_{i,t} \tag{7}$$

with $\nu_{i,t} \sim N(0, \sigma_i^2)$. The final set of time-varying parameters, the free elements of the matrix A_t, has equivalent dynamics:

$$\alpha_t = \alpha_{t-1} + \zeta_t \tag{8}$$

Finally, all the innovations are assumed to be jointly normally distributed with the following variance covariance matrix:

$$V = Var\left(\begin{bmatrix} \epsilon_t \\ \eta_t \\ \zeta_t \\ \nu_t \end{bmatrix}\right) = \begin{bmatrix} I_N & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix} \tag{9}$$

where Q, S and W are positive definite matrices. The restrictions on V are aimed at making the model as more parsimonious as possible given the already high dimensionality imposed.[14]

4.2 Priors

Priors choices are here outlined. Their calibration is mostly based on the first ten years of data, using observations from 1953:I to 1962:IV.

The initial condition for the slope parameters is normally distributed with mean and variance related to the OLS point estimates of a time invariant VAR estimated on the initial subsample,

[14] As Primiceri (2005) notes, the independence between blocks of innovations is anyway not a necessary restriction for the entire approach to work.

$$\theta_0 = N(\theta_{OLS}, 4 * V(\theta_{OLS})). \tag{10}$$

The diagonal elements in the matrix Σ_t are treated as states and are assumed to follow a geometric random walk. As for their initial values, the usual log-normal distribution is assumed

$$f(ln\sigma_{i0}) = N(ln\bar{\sigma}_i, 1) \tag{11}$$

where $\bar{\sigma}_i$ equals the square root of the i-th diagonal element of the lower triangular matrix derived from the Choleski decomposition of the presample residual variance. Cogley and Sargent (2005) and BG use a variance of 10 which implies a dramatic variability around the prior mean but they model the log variance rather than the log volatility as a state. In this respect, I follow Primiceri (2005).[15]

The prior distribution for the off-diagonal elements of the lower diagonal matrix A_t is assumed to be normal,

$$f(\alpha_0) = N(\bar{\alpha}, 10 * I_{n*(n-1)/2}) \tag{12}$$

where $\bar{\alpha}$ also is obtained from the Choleski decomposition of the presample residual variance, by imposing appropriate restrictions on the elements of the latter.

The model parameters, instead, are all supposed to have an inverse-Wishart distribution:

$$Q \sim IW(k_Q^2 * 40 * V(B_{OLS}), 40) \tag{13}$$
$$W \sim IW(k_W^2 * 40 * I_n, 40) \tag{14}$$
$$S_1 \sim IW(k_S^2 * 2 * 10 * I_1, 2) \tag{15}$$
$$S_2 \sim IW(k_S^2 * 3 * 10 * I_1, 3) \tag{16}$$
$$S_3 \sim IW(k_S^2 * 4 * 10 * I_1, 4) \tag{17}$$

[15]Furthermore, a unit variance for the log volatility seems reasonable for what such a value implies for the associated prior belief on the volatility itself. Recall that the latter can be equivalently considered as a log-normal random variable whose moments are severely affected by our choices about $ln(\sigma_{i0})$: unlike posterior medians, it will be surely reflected in the width of posterior bands.

where S_1, S_2 and S_3 denote the three blocks of S and $k_Q = k_W = 0.01$ while $k_S = 0.1$. These priors are diffuse or at most weekly informative. The structure of S outlined above refers to the case of a fourvariate VAR but it can be easily generalized.

4.3 The MCMC algorithm

The Gibbs sampling algorithm is used along the lines of Primiceri (2005) in order to simulate the posterior distribution of the hyperparameters and the states. Here follows a very coincise summary of all the steps:

- Initializing hyperparameters and states;
- Sampling the vector θ^T conditional on the data and all the remaining states and hyperparameters;
- Sampling the time-varying elements of the matrix A^T conditional on the data and all the remaining states and hyperparameters;
- Sampling the diagonal elements of the matrix Σ^T conditional on the data and all the remaining states and hyperparameters;
- Sampling the non-zero elements in V from their respective conditional posteriors

where x^t denotes the entire history of the vector (matrix) x up to time t, T being the sample length. Algorithms and simulations in each step are quite standard but any interested reader can refer to Primiceri (2005) for a detailed description of the posterior simulator. The stability constraint is imposed after drawing the vector θ^T and the reference is Cogley and Sargent (2005).

Under certain regularity conditions, by iterating the previous steps for a sufficient number of times one can treat the resulting draws as taken from the joint posterior distribution. In this paper, I use a burn-in period of 5000 iterations to converge to the ergodic distribution and retain every other draw of the subsequent 30,000 iterations in order to reduce the autocorrelation across draws. Hence, the inference is based on 15,000 draws. Convergence has been checked by means of the Geweke's CD statistic and the Inefficiency Factor.

5 Empirical evidence

I have estimated the model outlined in section (4) for both the bivariate (output growth and yield spread) and the fourvariate case (output growth, inflation, yield spread and the short term rate). The lag order has been set to p=2 consistently with most of the literature. As in BG and Cogley and Sargent (2005), a stability constraint is imposed on the system at each point in time. This is represented by an indicator function $I(\theta^T) = \prod_{s=1}^{T} I(\theta_s)$ that takes value one when the roots of the associated VAR polynomial are outside the unit circle.[16]

The estimates from the VAR are of little interest per se, especially because the focus is on multi-period forecastability. Anyway, some information can be obtained by glancing at the patterns exhibited from volatilities and parameters. Focusing on the bivariate case, Figure 2 confirms the presence of decreased residual variability for output growth from the second half of the eighties. For the yield spread there is a spike in the early eighties, instead, which coincides with the Volcker's "monetarist experiment".[17]

On the contrary, parameters do not move much, rather they appear to be flat over the entire sample (see Figures 3 and 4).[18] The yield spread affects output growth two period-ahead: the sign of the coefficient is positive as expected but its value has decreased over the last decade. On the other hand, GDP growth is negatively correlated with future yield spread, which also displays a lot of inertia.

Ultimately, the aim of the paper is to ascertain whether the predictive power of the yield spread is material, has changed and, finally, how. To this end, BG use spectral techniques to measure model fit and predictability at 1, 4 and 8 periods ahead.[19] I have, instead, followed

[16]The stability condition is actually checked by rewriting the VAR specification in (1) in its companion form and verifying all the eigenvalues are less than one in absolute value. Cogley and Sargent (2005) explain how such procedure works and how the relative priors and posteriors are modified with respect to the unconstrained version. Alternatively, Koop and Potter (2008) suggest to impose the reflecting barrier and use the single-move rather than the multi-move sampling of the vector θ^T. The former is decidedly more inefficent but a rejection of a single θ^t does not imply the rejection of the entire θ^T, thus being preferable to multi-move sampling in such a circumstance.

[17]The patterns for both output growth and yield spread are confirmed in a fourvariate VAR that includes inflation and the short rate as well (Figure 5).

[18]I have found this evidence to be robust to various changes in prior's hyperparameters. For example, even ignoring the reflecting barrier as in Primiceri (2005) delivers the same results.

[19]The concepts of fit and predictability inevitably coincide here. As for the technical aspects, inference on predictability is entirely based on a multivariate definition of persistence measured by the normalized specturm

Cogley, Primiceri and Sargent (*forthcoming*) (henceforth CPS) who look at a different measure of persistence that is claimed to be more precise than the previous one. Such a statistic is still the ratio of conditional to unconditional variance but the two terms are related to the concept of total variation of future variables at a given horizon due to past and future shocks, respectively.[20] In more detail, consider the companion form of the VAR

$$z_{t+1} = \mu_t + \Gamma_t z_t + \epsilon_{z,t+1} \tag{18}$$

where the vector z_t contains current and lagged values of Y_t, the vector μ_t contains the VAR intercepts and the matrix Γ_t instead includes the autoregressive parameters. CPS use this representation for multi-step forecasting assuming that all the parameters remain constant over time from that moment onwards.[21] Hence the j-period ahead forecast error variance is approximated by

$$var_t(\widehat{z}_{t+1}) \approx \sum_{h=0}^{j-1} (\Gamma_t^h) var(\epsilon_{z,t+1}) (\Gamma_t^h)' \tag{19}$$

while the unconditional variance of $\widehat{z}_{t+1}$ is obtained as the limit of the conditional variance as the forecast horizon j increases,

$$var(\widehat{z}_{t+1}) \approx \sum_{h=0}^{+\infty} (\Gamma_t^h) var(\epsilon_{z,t+1}) (\Gamma_t^h)'. \tag{20}$$

Under the anticipated-utility approximation, this is also the unconditional variance of $\widehat{z}_{t+s}$ for $s > 1$. The final measure of predictability is obtained as one minus the fraction of the total variation due to future shocks or, given that future shocks account for the forecast error, as one minus the ratio of the conditional variance to the unconditional variance,

at frequency zero as typically done in the literature. The only drawback is that stability has to be imposed in the system at each time t in order to apply frequency domain methods.

[20] Even Cogley, Primiceri and Sargent (*forthcoming*) anyway use the reflecting barrier on the parameter vector. Remember that incercepts are left free to wander, though, as they do not influence the stability of the VAR representation.

[21] They note that such an approximation is common in the literature on bounded rationality and learning. Furthermore, they report in other papers of theirs that it approximates well the mean of Bayesian predictive densities.

$$R^2_{j,t} = 1 - \frac{var_t(e_\pi \widehat{z}_{t+1})}{var(e_\pi \widehat{z}_{t+1})} \approx 1 - \frac{e_\pi[\sum_{h=0}^{j-1}(\Gamma_t^h) var(\epsilon_{z,t+1})(\Gamma_t^h)']e'_\pi}{e_\pi[\sum_{h=0}^{+\infty}(\Gamma_t^h) var(\epsilon_{z,t+1})(\Gamma_t^h)']e'_\pi} \tag{21}$$

where e_π is a selection vector. This $R^2_{j,t}$ is analogous to the R^2 statistic for j-step ahead forecasts and must lie between zero and one. The focus is here only on the predictability of 4 and 8 quarters ahead output growth as these are the horizons of major interest to policymakers.

In the following, the analysis is based on the fourvariate specification.[22] Figure 6 shows the median and the interquartile range of the four periods ahead $R^2_{j,t}$ for output growth. Predictability reaches its peak at about 60% in the very early eighties, then declines and stabilizes at about 20% in the following decade. Only from 2000 on it increases again with occasional sharp hikes, among which the first quarter of 2009.

These graphs anyway draw a picture based on the contribution of all the variables included in the multivariate specification while the aim is to gauge the marginal contribution of the yield spread. Following BG, this is done by computing for each horizon of interest the marginal $R^2_{j,t}$ as the difference between the $R^2_{j,t}$'s of the fourvariate VAR and the $R^2_{j,t}$ of the trivariate VAR without the yield spread. It has to be said at the outset that such quantity is inherently stochastic and some values are drawn that are negative. Negative values are, in any case, to be interpreted as sign of unpredictability. At $h = 4$ (see Figure 8), the term spread appears to add some forecasting power in the eighties. After falling in the nineties its predictive content starts increasing again around 2001 and 2002 but then vanishes lately.[23] At $h = 8$ (see Figure 9) almost a flat line, close to zero, is instead obtained suggesting that there is no additional information.

It can be checked to what extent the amount of predictability, as captured by the marginal

[22] The reason is twofold. Firstly, as argued by BG, it is important to investigate "whether the spread contains information which is not already encoded in other macroeconomic variables, first and foremost measures of the monetary policy stance..", while the extant literature has not gone beyond the bivariate case. Secondly, jumping to the discussion about the marginal predictive power, Figures 11 and 12 show a dramatic amount of uncertainty that obscures the clear pattern exhibited by the median of the posterior distributions of the marginal $R^2_{j,t}$. This pattern is also evident from Figures 13 and 14 that plot the joint posterior distribution of the marginal $R^2_{j,t}$ at selected points in time.

[23] For the sake of clarity, these statements are based on the median of the posterior distribution of each marginal $R^2_{j,t}$. Please note that as made clear also in BG, it is very unlikely to reject the hypothesis that such a statistic is insignificant in almost all the sample. Nonetheless, it remain interesting to further investigate on the episodes in which predictability seems to be stronger.

R_t^2, has changed over time. The assessment is based on scatter plots of the distribution of the four quarters ahead marginal R^2 at selected points in time. In particular I have compared three-quarter moving averages computed at the peak (1981:I), at a moment of constant low predictability (1997:II) and finally in the last "available" quarter (2009:I), respectively. In the subplots of Figure 10 most of the points lie under the 45-degree line which gives further support to the view that predictability was higher during the early eighties.[24]

The evidence so far is in favour of that strand of literature stating that the yield spread does not convey additional information to predict future fluctuations in output once the short term rate is taken into account. There are times, on the other hand, when the predictive content evidently increases and it would be interesting to investigate in more detail why and what drives the evidence. In this respect BG note a striking coincidence with periods in which the uncertainty about the future monetary policy stance is high but no formal analysis is there attempted. Others, instead, point to the role of inflation expectations that are a fundamental component of term *premia* and then look at the variance of long term rates.

A similar, yet not identical, route is here suggested. On the premise that uncertainty (or equivalently risk aversion) may impact on the relationships linking economic variables, it could be tentatively explored whether a proxy like the default spread helps explain the (*in sample*) perfomance of the yield spread, for instance by computing a simple measure of correlation between the default spread (BAA-AAA) and the marginal R_t^2. A more rigorous approach would entail augmenting the previous fourvariate VAR with this new variable. While this is appealing on principle, the practical implementation raises problems in terms of dimensionality of the entire system making the estimation particularly intensive from a computational point of view.

6 Conclusions

In the present work I investigate the predictive power of the yield spread for a measure of economic activity, namely GDP growth. Even in the absence of a dominant theoretical framework,

[24]As remarked above when commenting on Figure 11, each marginal R_t^2 sometimes happens to take negative values as a consequence of its stochastic nature. This is of course reflected also in the scatter plots of Figure 10.

empirical papers addressing this topic abound. This work adds to the strand of the literature focusing particularly on the time varying nature of this predictive relationships.

A wide consensus points to a breakdown started in coincidence with the "Great Moderation" period but the last recession prompts to a reconsideration or, at the very least, to a new and updated assessment. This is done within a multivariate approach that allows for smooth variation in parameters, conditional variances and correlations. The evidence suggests that the predictive power still remains low lately and that the peak has been reached in the early eighties. Overall the yield spread seems not to convey a significant signal once inflation and the short term interest rate are also considered.

Furthermore, an interesting question is whether the inclusion of the default premium (the difference between BAA and AAA rated bond yields) affects the performance of the yield spread. A first step in this direction could be to compute the plain correlation between the former and the measured marginal contribution of the latter. A more detailed analysis is anyway left for future research.

Finally, it has to be said that a proper way of dealing with all the aspects related to forecasting entails embracing a real time approach, whereas all the evidence here (and BG) is conditioned on the full sample.[25] However, a recursive approach of the same kind as explained in section 4.2 would impose enormous costs in terms of computational time as to make the analysis unfeasible.

[25] Some methods have been recently developed in Bayesian statistics for this purpose: they are a class of Sequential Monte Carlo (SMC) methods.

References

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

[2] Benati L. and Goodhart C., "Investigating time-variation in the Marginal Predictive Power of the Yield Spread", *Journal of Economic Dynamics & Control*, 2008, N.32, 1236-1272.

[3] Benati L. and Surico P., "VAR Analysis and the Great Moderation", *The American Economic Review*, 2009, N.99, pp. 1636-52.

[4] Bordo M.D. and Haubrich J.G., "The Yield Curve, Recessions and the Credibility of the Monetary Regime: Long Run Evidence, 1875-1997", *NBER Working Paper*, 2004, N.10431.

[5] Bordo M.D. and Haubrich J.G., "The Yield Curve as a Predictor of Growth: Long Run Evidence, 1875-1997", *Review of Economics and Statistics*, 2008, N.90, 182-185.

[6] Cogley T., "How Fast Can the New Economy Grow? A Bayesian Analysis of the Evolution of Trend Growth", *Journal of Macroeconomics* ,2005, N.27, 179-207.

[7] Cogley T., Primiceri G. and Sargent T.J., "Inflation-Gap Persistence in the U.S. ", *American Economic Journal: Macroeconomics*, forthcoming.

[8] Cogley T. and Sargent T.J., "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.", *Review of Economic Dynamics*, 2005, N.8, 262-302.

[9] D'Agostino A., Giannone D. and Surico P., "(Un)Predictability and Macroeconomic Stability", *CEPR Discussion Paper*, 2007, N.6594.

[10] De Pace P., "GDP Growth Predictions through the Yield Spread Time-Variation and Structural Breaks", *Manuscript*, 2009.

[11] Estrella A., "Why Does the Yield Curve Predict Output and Inflation?", *The Economic Journal*, 2004, N.115, 722-744.

[12] Estrella A., Rodrigues A. and Schich S., "How stable is the Predictive Power of the Yield Curve? Evidence from Germany and the United States", *The Review of Economics and Statistics*, 2003, N.85, 629-644.

[13] Favero C., Kaminska I. and Soderstrom U., "The Predictive Power of the Yield Spread: Further Evidence and a Structural Interpretation", *CEPR Discussion Paper*, 2005, N. 4910.

[14] Giacomini R. and Rossi B., "How Stable is the Forecasting Performance of the Yield Curve for Output Growth?", *Oxford Bulletin of Economic and Statistics*, 2006, N.68, 783-795.

[15] Giannone D., Lenza M. and Reichlin L., "Explaining the Great Moderation: It Is Not The Shocks", *ECB Working Paper*, 2008, N.865.

[16] Hamilton J. and Kim D.H., "A Reexamination of the Predictability of Economic Activity Using the Yield Spread", *Journal of Money, Credit and Banking*, 2002, N.34, 340-360.

[17] Harvey C., "The Real Term Structure and Output Growth", *Journal of Financial Economics*, 1988, N.22, 305333.

[18] Kim D. and Wright J., "An Arbitrage Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yileds and Distant-Horizon Forward Rates ", *Finance and Economics Discussion Series*, Board of Governors of the Federal Reserve System, 2005.

[19] Koop G., Jochmann M. and Strachan R.,"Bayesian Forecasting using Stochastic Search Variable Selection in a VAR Subject to Breaks", *International Journal of Forecasting*, forthcoming.

[20] Koop G. and Potter S., "Time Varying VARs with Inequality Restrictions", *Mimeo*, 2008.

[21] Mumtaz H. and Surico P., "The Transmission of International Shocks: A Factor-Augmented VAR Approach", 2009, *Journal of Money, Credit & Banking*, 2009, N.41, 71-100.

[22] Rosenberg J.V and Maurer S., "Signal or Noise? Implications of the Term Premium for Recession Forecasting", *FRBNY Economic Policy Review*, 2008.

[23] Rudebusch G., Sack B. and Swanson E., “Macroeconomic Implications of Changes in the Term Premium”, *Federal Reserve Bank of St. Louis Review*, 2007, N.89, 241-269.

[24] Wheelock D.C. and Wohar M.E., “Can the Term Spread Predict Output Growth and REcessions? A Survey of the Literature”, *Federal Reserve Bank of St. Louis Review*, 2009, N.91, 419-440.

[25] Zhu X. and Rahman S.,“A Regime Switching Macro-Finance Model of the Term Structure”, *Economic Growth centre Working Paper Series*, 2009.

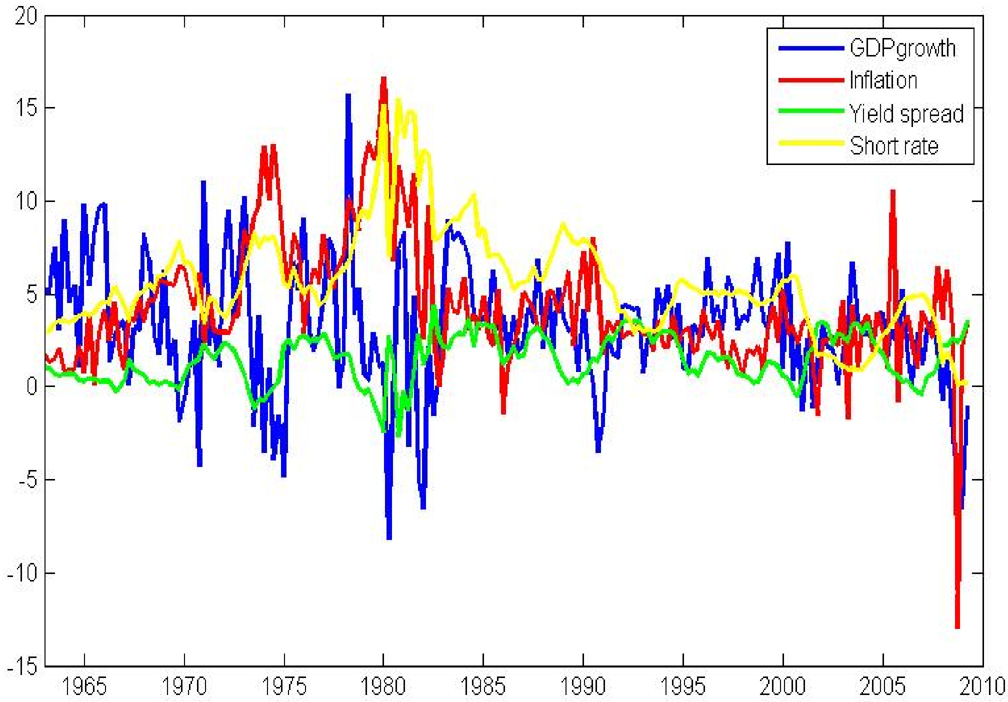

Figure 1: Plot of the series (1963:I-2009:II)

Table 1: Unconditional Correlation matrix (1963:I-2009:II)

	GDP growth	*Inflation*	*Yield spread*	*Short rate*
GDP growth	1.00			
Inflation	−0.0818	1.00		
Yield spread	0.0205	−0.3776	1.00	
Short rate	−0.0277	0.6075	−0.4198	1.00

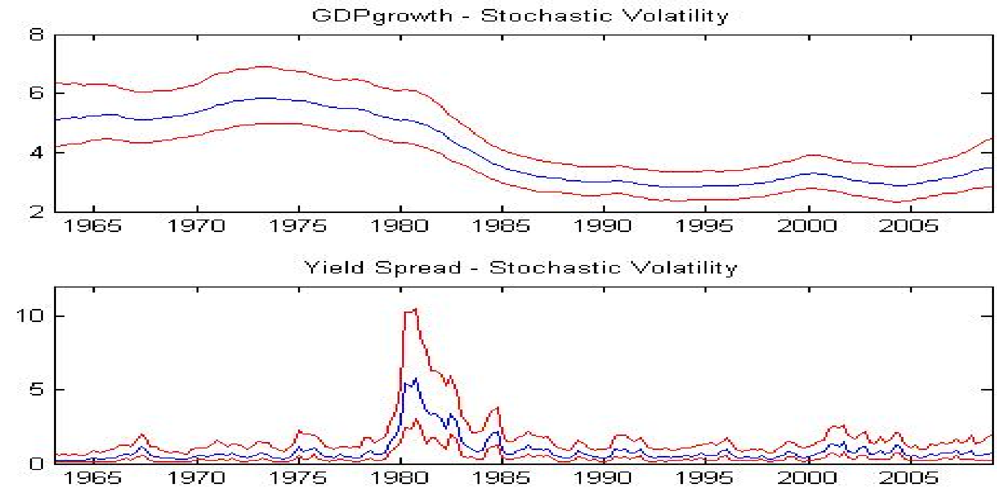

Figure 2: Stochastic volatilities - bivariate VAR

Notes: The graph shows the median, the 16th and the 84th percentiles of the posterior distribution of the stochastic volatilies in the bivariate VAR.

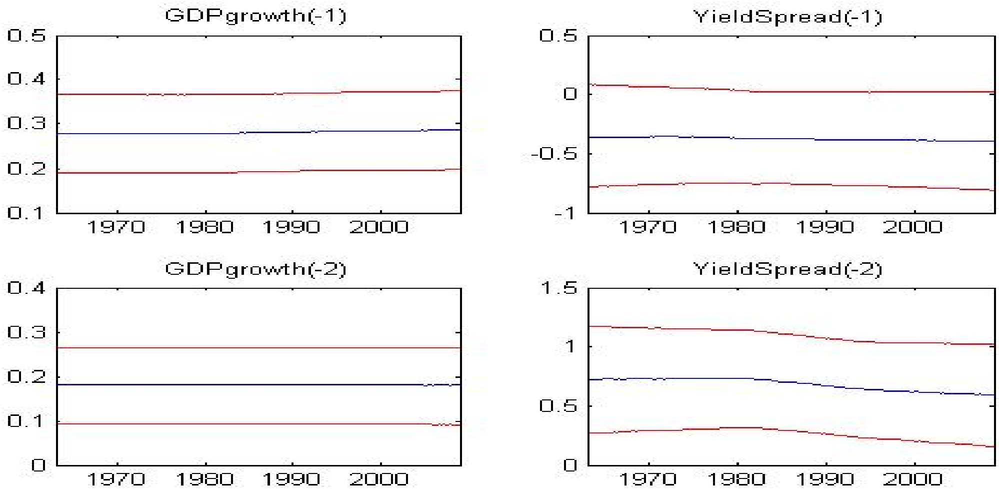

Figure 3: Time varying parameters - bivariate VAR

Notes: The graphs shows the median, the 16th and the 84th percentiles of the posterior distribution of the time-varying parameters in the OUTPUT GROWTH equation.

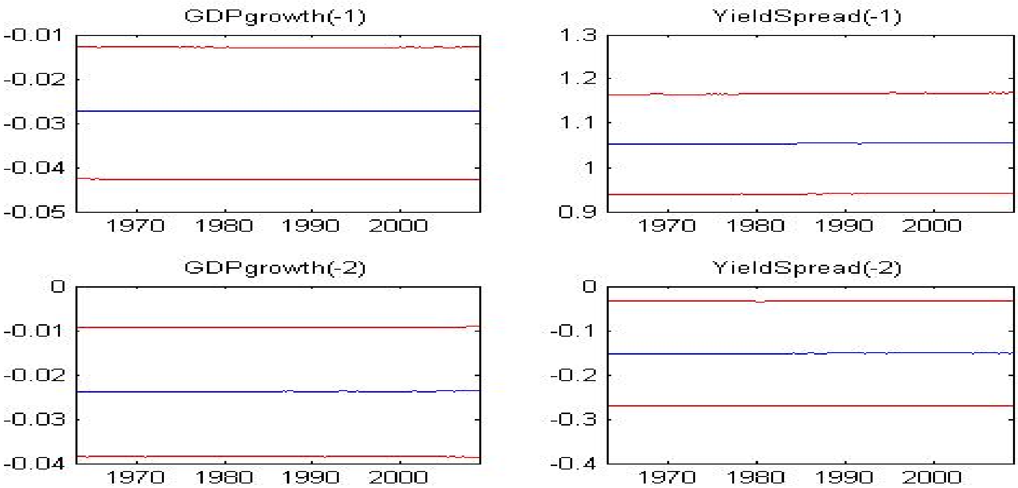

Figure 4: Time varying parameters - bivariate VAR

Notes: The graphs shows the median, the 16th and the 84th percentiles of the posterior distribution of the time-varying parameters in the YIELD SPREAD equation.

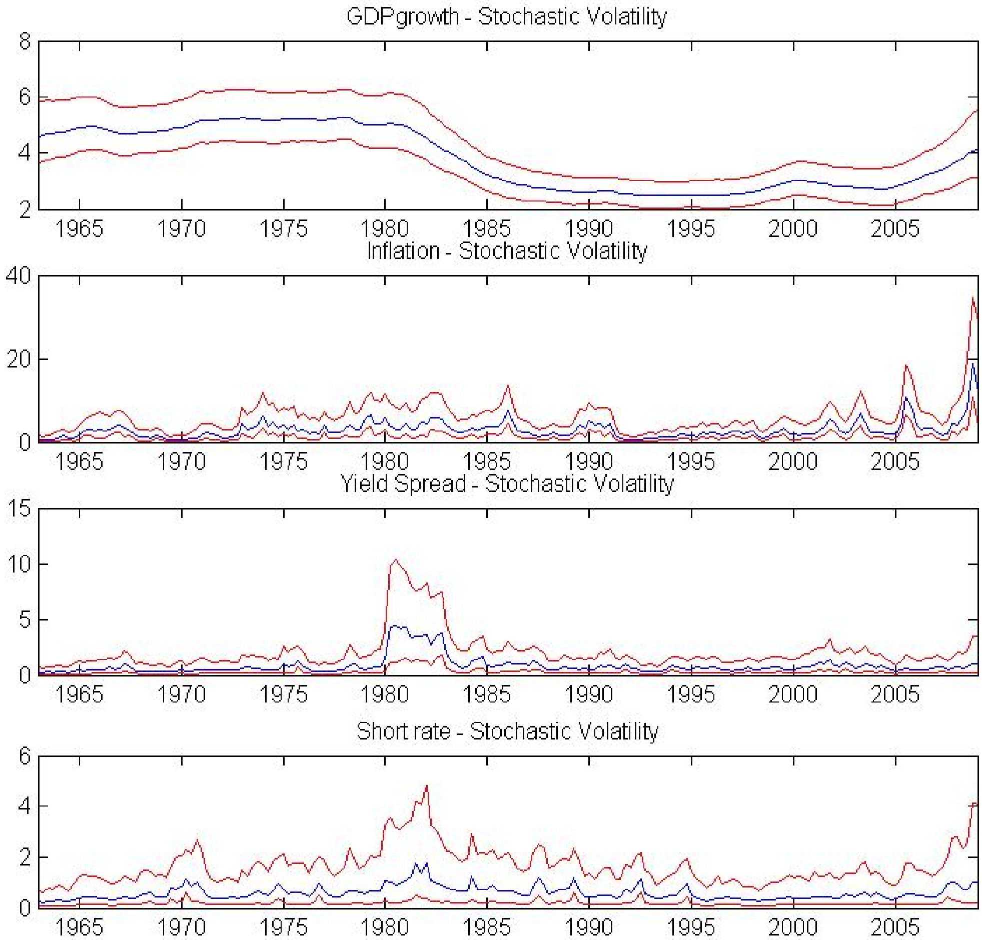

Figure 5: Stochastic volatilities - fourvariate VAR

Notes: The graph shows the median, the 16th and the 84th percentiles of the posterior distribution of the stochastic volatilies in the fourvariate VAR.

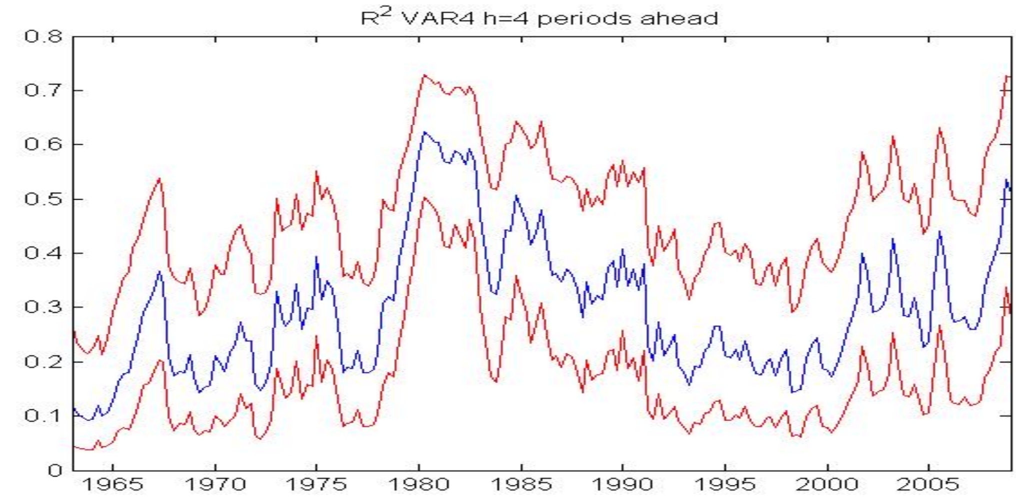

Figure 6: Overall R^2 at $h = 4$ periods ahead - fourvariate VAR

Notes: The graph shows the median and the interquartile range of the output growth's overall R_t^2 computed from the fourvariate model in section 4.

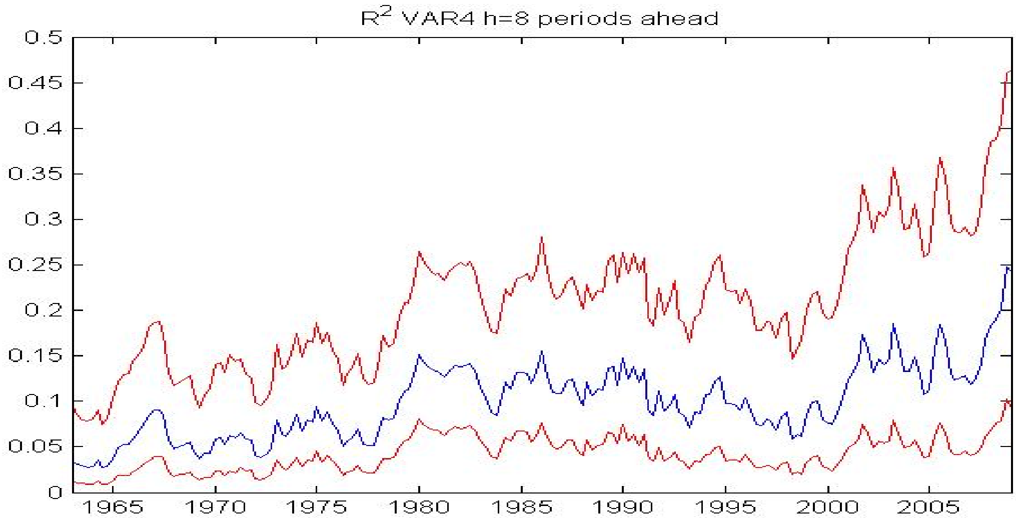

Figure 7: Overall R^2 at $h = 8$ periods ahead - fourvariate VAR

Notes: The graph shows the median and the interquartile range of the output growth's overall R_t^2 computed from the fourvariate model in section 4.

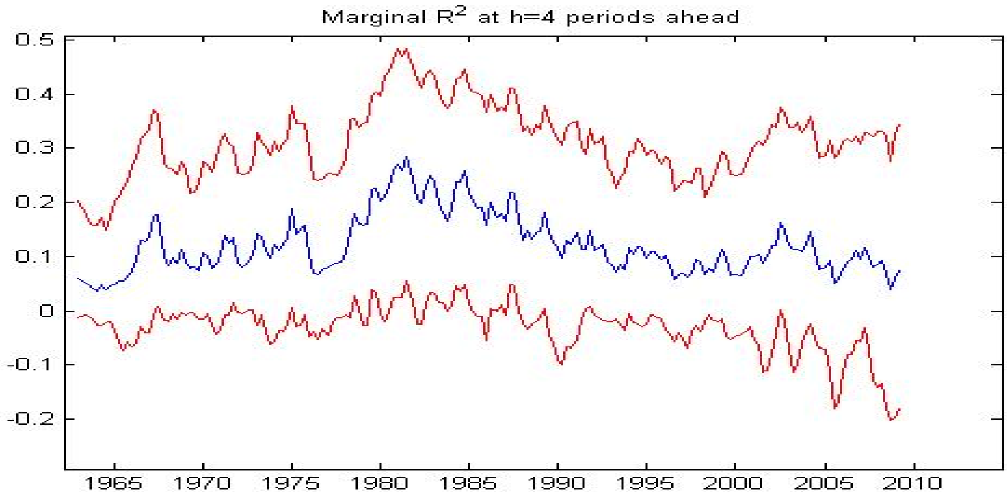

Figure 8: Marginal R^2 at $h = 4$ periods ahead - fourvariate VAR

Notes: The graph shows the median and the interquartile range of the marginal R_t^2 computed as the difference between the fourvariate and the trivariate models in section 4 to assess the marginal predictive content of the yield spread.

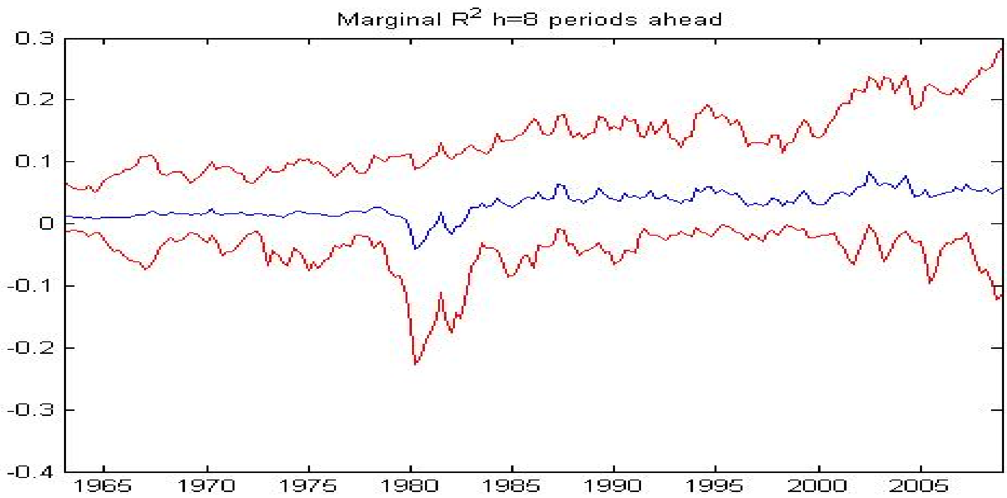

Figure 9: Marginal R^2 at $h = 8$ periods ahead - fourvariate VAR

Notes: The graph shows the median and the interquartile range of the marginal R_t^2 computed as the difference between the fourvariate and the trivariate models in section 4 to assess the marginal predictive content of the yield spread.

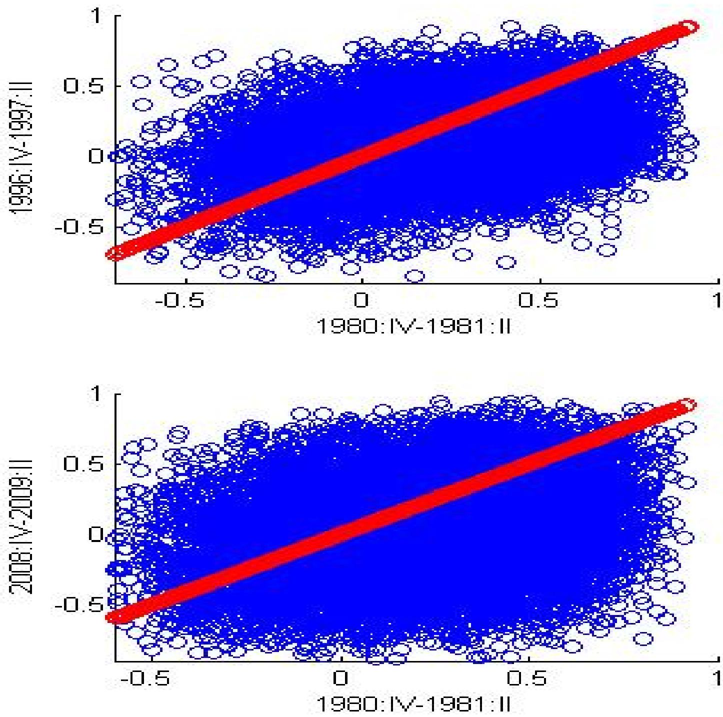

Figure 10: Joint distributions of the marginal R^2, $h = 4$ periods ahead - bivariate VAR

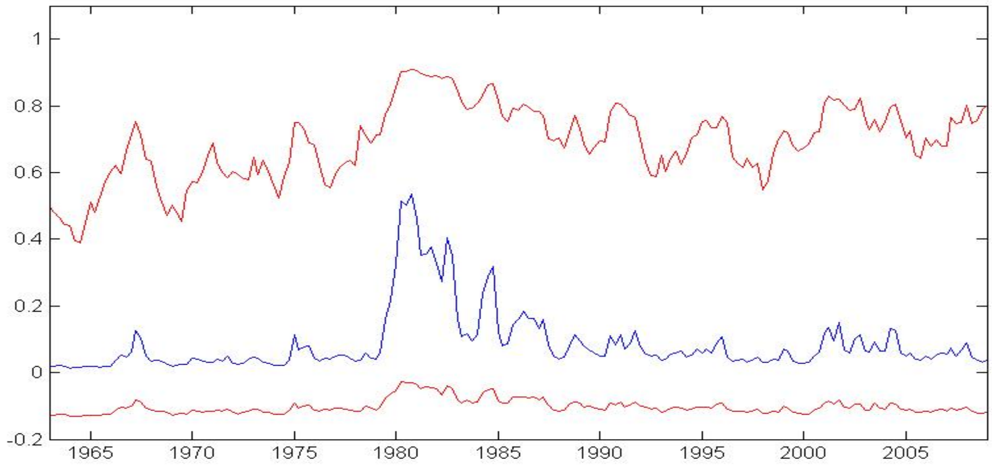

Figure 11: Marginal R^2 at $h = 4$ periods ahead - bivariate VAR

Notes: The graph shows the median and the interquartile range of the marginal R_t^2 computed as the difference between the bivariate and the univariate models in section 4 to assess the marginal predictive content of the yield spread.

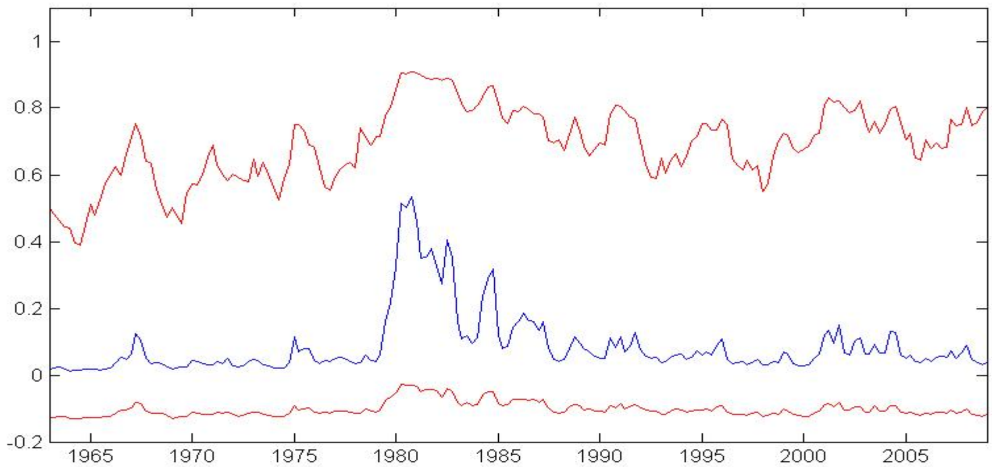

Figure 12: Marginal R^2 at $h = 8$ periods ahead - bivariate VAR

Notes: The graph shows the median and the interquartile range of the marginal R_t^2 computed as the difference between the bivariate and the univariate models in section 4 to assess the marginal predictive content of the yield spread.

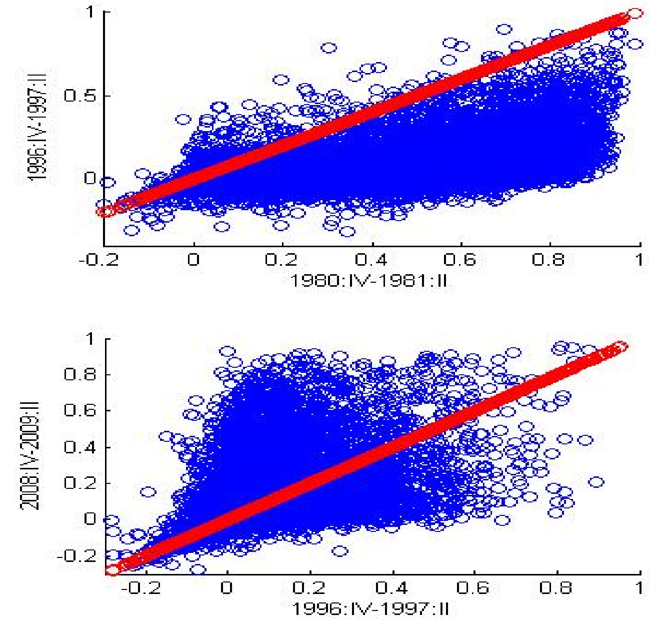

Figure 13: Joint distributions of the marginal R^2, $h = 4$ periods ahead - bivariate VAR

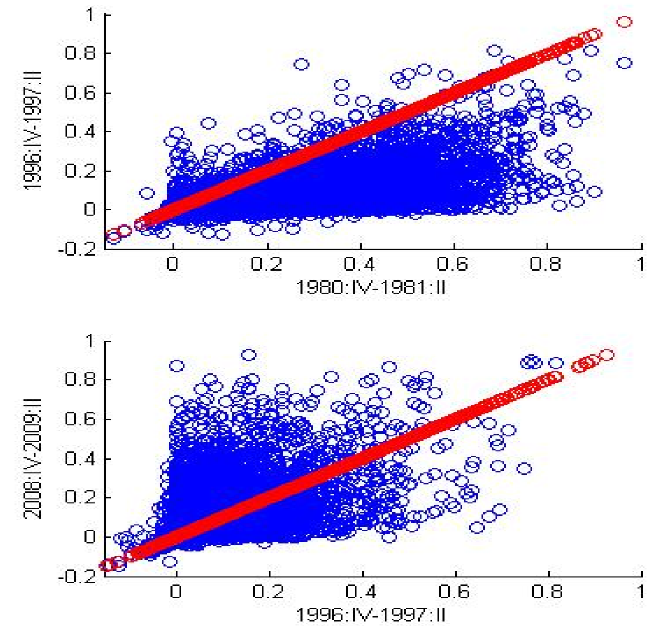

Figure 14: Joint distributions of the marginal R^2, $h = 8$ periods ahead - bivariate VAR