Instituto Nacional de Matemática Pura e Aplicada

Doctoral Thesis

Essays on Macro-Finance: Identification, Estimation and Forecasting of Term Structure Models with Macro Factors and Default Risk

Marco S. Matsumura

Advisor: Aloisio Araujo

September 2008

Thesis committee:

Aloisio Araujo, IMPA & FGV

Caio Ibsen, FGV

Hélio Migon, UFRJ

J. Valentim Vicente, BACEN & IBMEC

Márcio Nakane, USP & BACEN

Dedicado à minha família.

Agradecimentos

Sou especialmente grato a toda comunidade do IMPA, na qualidade de seus professores, funcionários e colegas, pela oportunidade única na vida de participar de um ambiente tão rico, criativo e estimulante para a pesquisa e muito além, nesses anos cruciais para minha formação, e pelo apoio ágil e ‡exível da instituição nos momentos importantes. Agradeço ao Professor Aloísio Araújo por sua orientação e grandes ensinamentos, e por toda liberdade de pesquisa que pude, talvez imerecidamente, desfrutar. Sou grato aos meus Professores de Economia Matemática, Wilfredo Leiva, Paulo Klinger e Humberto Moreira, que me ensinaram os primeiros passos numa área que me era então completamente nova. Agradeço ao Professor Claudio Landim, que num curso de verão há tantos foi quem primeiro incentivou a ingressar no IMPA, e que considero o começo de todo o processo que gerou este trabalho. Agradeço também a todos os Professores e Amigos e Colegas de turma, com os quais sempre procurei aprender o máximo que pude, mas a quem certamente devo mais do que jamais poderei retribuir. Devo ao IPEA a liberdade com que pude escolher e manter minha linha de pesquisa desde minha entrada na instituição, e às ótimas condições de trabalho. Agradeço à banca de doutorado, que muito interagiu e muito contribuiu nos estágios …nais da preparação deste texto. Fundamentais também para a realização deste trabalho foram os apoios das agências de fomento CNPq e Capes, dos quais fui continuamente bolsista desde meu primeiro ano da graduação.

Contents

1. General introduction

This work propose a number of improvements in the area of macro-…nance models, which consists of combinations of term structure and macroeconometric models.

There are multiple applications of term structure models with di¤erent ends. Some of which are given in the following: 1) …nancial market practitioners need models for interest rate and credit risk derivatives; 2) in order for the Central Bank to conduct the monetary policy and monitor the yield curve, it need to know how the curve is related to macroeconomic indicators; 3) the Treasury continuously demands assessments of current and future interest rates to manage the emission and maintenance of the stock of public debt.

We assess the impact of macro factors on the yield curve of emerging countries, where data is relatively scarce and volatile as compared to developed countries'. Also, frequent changes of regime, crisis episodes and defaults limit the extent of the existing historical series.

In those markets, the Monetary Authorities may respond to other factors besides the expected in‡ation and the output gap. The exchange rate, for instance, is an important variable.

Moreover, the relevance of explicitly including macro factors into the term structure models may di¤er when dealing with developed or with emerging markets. The …xed income markets are huge and very sophisticated markets in which assets are traded in almost continuous-time. Do bond investors already process all the markets information e¢ ciently?

In contrast, macroeconomic indicators have a low frequency nature as they must be collected and processed by institutions; they are not real-time information. For this reason, traditionally the macro-…nance models are discrete-time, and estimated with monthly or quarterly data. Ang and Piazzesi (2003) estimated their model with a quarterly series containing roughly 50 years of data.

On the other hand, the limitations of emerging market samples may require the use of daily data. In this case, continuous-time models constitute a natural choice.

In our …rst article, we analyze the relation between the Brazilian domestic term structure and macroeconomic variables.

We construct a novel dataset consisting of daily samples of 1) DI x Pre interest rate swaps, traded on the BM&F, proxy for zero-coupon constant maturity term structure, 2) INPC x DI swaps (available since 2002), which provide a measure of daily in‡ation, and 3) the US Dollar/ Brazilian Real exchange rate.

We show that the inclusion of macro factors signi…cantly improves the goodness-of-…t of the term structure models by comparing yields alone and macro-augmented a¢ ne models.

Continuous and discrete-time Gaussian a¢ ne models are considered. The measurement errors are treated with either Kalman …lter or Chen-Scott inversion. Three di¤erent Taylor rules and two di¤erent lag sizes are considered, and the parameters are inferred through maximum likelihood or Monte Carlo Markov chain (MCMC).

Among our empirical …ndings, we remark that macro factors explain about 40% of the movements of the yield curve, and that imposing restrictions on the speci…cation of the model - such as macro-to-yield dynamics or standard Taylor rule - signi…cantly modi…es the response of the yield curve, and should thus be carefully considered.

In chapter two, we extend the models by adding default risk and address the problem of …nding the factors behind the movements of emerging market sovereign interest rates. The models are estimated again with daily data, allowing for the interaction between macro variables and credit spreads.

The emerging markets' Central Banks do not have direct in‡uence on sovereign interest rates, which are traded on international markets. Thus, contrary to the domestic case, there are no obvious candidates as the most important macro factors in‡uencing the external rates.

We calculate default probabilities implied from the estimated model and the impact of macro shocks on those probabilities.

Our empirical results show that, given our tested variables and horizon, the VIX (a volatility index calculated using S&P 500 option prices) is the most important macro factor a¤ecting short term bonds and default probabilities, while the Fed Fund is the most important factor behind long term default probabilities. Regarding tested domestic factors, only the slope of the domestic yield curve showed relevant e¤ect.

Before estimating the models, they must …rst be identi…ed. In fact, we show that the likelihood function of the Gaussian a¢ ne models with macro factors and default using Chen-Scott inversion are invariant to certain operators.

The question of identi…cation is not completely discussed by the literature. Our contributions to this topic are the subject of our third article, where we extend the method of Dai and Singleton (2002) to term structure models with macro factors.

There many possible identi…cations, but we prove that the choice of identi…cation does not a¤ect the response of the yield curve or of the macro factors to state variable shocks. However, it does a¤ect the latent factor response.

Next, in article four, we focus our attention on the out-of-sample predictive capacity of a number of term structure models. The models are estimated with Brazilian and U.S. daily samples, so that we obtain results that contrast speci…c characteristics of emerging and developed markets data.

In particular, we compare the a¢ ne and the Nelson-Siegel models, including the Diebold-Li 2-step model, which list among the most popular in the …nancial and econometric literature.

We test whether macro-…nance models outperforms yields alone models. For Brazil, we consider a version with the IBovespa (Sao Paulo Stock Exchange Index) besides versions with expected in‡ation and output gap.

We show that for U.S. yields-only models already present good performance, and that the addition of macro factors does not improve the predictions.

For Brazilian data, both the yields only or macro-…nance models present low performance. However, the IBovespa signi…cantly contributes to the forecasting performance of the models.

On overall, the best model was the unrestricted dynamic Nelson-Siegel model.

In our …nal article, chapter 5, we further examine the properties of the macro-…nance models proposed in article 4. However, instead of forecasting, our focus turn to calculate impulse response functions and variance decompositions so as to analyze the nature of the impact of the di¤erent indicators on the yield curves of both markets.

We discuss a novel model combining monthly macro data and daily term structure data (a "pooled" model), which provides more accurate results than purely monthly models, specially for the case of Brazil.

We also propose a simple extension that incorporates change of regime. It specifically tries to capture possible changes of monetary policy. It is estimated with Greenspan and Bernanke periods, and presidents Lula and Cardoso's periods.

All models in chapter 4 and 5 are estimated via MCMC. In the appendix we perform a series of simulation exercises were we show that our MCMC algorithm correctly recovers the true parameters, and that the chains generated by the Gibbs sampling and Metropolis-Hastings converge under the Gelman-Rubin diagnostics.

CHAPTER 1

Assessing Macro In‡uence on Brazilian Yield Curve with A¢ne Models

1. Introduction

The term structure of interest rates synthesizes agents' perceptions about the future state of the economy. The interaction between that perception and macroeconomic variables is an important element for consideration by the monetary authorities (MA) for policy decisions and by market participants for forecasting. Ang and Piazzesi (2003), A&P, discuss that interaction proposing a model that combines ideas from the …nancial and the macroeconometric literature.

In the …nancial literature, the a¢ne term structure models (Du¢e and Kan, 1996), constitute a popular class of models, in which the yield and the risk premiums are modeled in continuous time as a¢ne functions of unobserved state variables. However, since standard a¢ne models do not contain macroeconomic variables, they cannot be directly related to the yield curve or latent factors.

Macroeconometric models usually analyze the e¤ect of observed variables on the yield curve, and model the dynamics of the rates and of the e¤ects of …nancial and macro shocks. But they do not take into account the no-arbitrage restrictions among the rates of the diverse maturities, which can potentially lead to an over-parameterization of the model and a reduction of its forecasting capacity.

A&P's model incorporates observed macro variables - the output gap and in‡ation - into a discrete-time a¢ne model and a MA reaction function to nominal shocks - the Taylor rule - to study the relation between the economic cycle and the yield curve. This is an extension of small-scale macro models such as that of Svensson (1997). In this way, some of the exogenous shocks on the state factors and their e¤ects on the yield curve become identi…able in a model with no arbitrage restrictions among the yield maturities.

By including macro variables, the task of the inference of the parameters becomes more di¢cult, due to the nonlinear character of the model. This fact motivated us and Ang et al. (2005) to use the Monte Carlo Markov chain (MCMC) algorithm, a Bayesian approach, which is less vulnerable to dimensional issues than maximum likelihood.

Our work addresses the Brazilian market, which requires di¤erent choices of frequency and macro factors. In A&P, the unity of time is the quarter, the frequency with which the output gap is measured. This is not possible for analyzing the Brazilian economy.

The behavior of emerging countries' …nancial markets can be distinguished from that of developed countries' markets by the lower liquidity, shorter term structure (less than 3 years until recently), more interventions that result in changes of regime and of rules of operation, existence of credit risk in public debt, greater vulnerability associated with volatile exchange rates, and limited data availability.

Also, up to 1994 the Brazilian economy experienced a long period of high inflation. Inflation rates as high as 90% a month occurred during this period.

However, even in this environment, the Central Bank was able to preserve the credibility of the domestic currency as a denomination of public debt and means of legal payment, thanks to an ample indexation system that included the Brazilian currency and the U.S. dollar exchange rate as a reference.

The Real Plan (including introduction of the present currency, the Real) was implemented in 1994 as a regime that pegged the domestic currency to the U.S. dollar until 1998, drastically reducing the inflation rates that prevailed at the time. In January 1999, this fixed exchange rate regime collapsed after a series of speculative attacks.

After that, the monetary authorities decided to adopt a free exchange rate combined with inflation targeting. The Central Bank issued local currency bonds linked to the dollar exchange rate index. A substantial fraction of the public debt was indexed to the dollar in this way. Together, these facts attest the importance of the exchange rate for the Brazilian economy.

One of the tools for monitoring the inflation targeting regime was developed by Bogdanski et al (2000). It consists of a macro model similar to that of Svensson (1997), extended with the effect of currency devaluations on the equation that determines the inflation rate. This modification was inspired by an extensive literature on emerging countries (see Fraga et al., 2003). Other articles produced by the Brazilian Central Bank (Almeida et al., 2003, Fachada, 2001, Freitas and Muinhos, 2001, Goldfajn and Werlang, 2000, Minella et al., 2003, Muinhos and Alves, 2003, Rodrigues et al., 2000) emphasize the importance of the exchange rate in price formation.

Backed by that evidence, we choose the inflation and exchange rate as our macro factors. And since our sample starts only after the alteration of the monetary regime in 1999, with the adoption of the inflation targeting, we use daily data. The unit of time used in A&P, the quarter (the frequency with which the output gap is measured), is not feasible for us.

Also, the sample size does not allow analysis of the iteration between real variables and the yield curve, inflation and exchange rate. Instead, we ignore the real variables and consider a reduced-form macro model with daily data.

A source of zero-coupon constant maturity rates of Brazilian domestic bonds is the BM&F Bovespa, the São Paulo Stock Exchange, which trades various types of interest rate swaps. Two of them are used here, the DIxPre and the INPCxDI. The first swap trades floating for fixed interest rates, and the second swap yields the floating rate on one side, and the expected consumer price index rates (INPC) on the other. The first swap permits the construction of the yield curve, and both together can be used to extract the expected inflation for certain horizons. The trading volumes of these two swaps are on average about U.S.$30 million each.

By using daily data, we focus on the financial, high frequency, aspects of the relation among interest rates, the exchange rate and expected inflation, as opposed to macroeconomic, low frequency aspects and larger horizons.

Due to the chosen unit of time, our model omits the output gap, which indicates the state of the economic cycle. However, the effect of the output gap is not absent:

since the yield curve summarizes the state of the economy, the latent factors contain information from omitted variables, and thus can capture economic cycle e¤ects.

One version of our model, (C), follows the tradition of a large body of …nancial literature that uses daily samples, speci…es the model in continuous-time and estimates it via maximum likelihood. Another version, (D), follows the econometric tradition, specifying a discrete-time model, and is estimated through MCMC.

It is not immediately evident how to choose the most adequate speci…cation and method of inference. Other technical questions, such as the forecasting horizon and the de…nition of the latent variable (discussed later), have to be decided by empirical testing. Thus, di¤erent speci…cations are estimated. We discuss the goodness-of-…t and impulse response functions in such a way that the two versions can be compared.

Many authors emphasize the importance of incorporating macro variables to …nancial models. For example, Diebold et al. (2005) remarks that pure a¢ ne models add little insight into the nature of the underlying economic forces driving the yield curve movements. Adding macro factors would shed light to the fundamental determinants of the interest rates. They point out the importance of the short rate as a fundamental building block to price all the bonds and as a policy instrument under direct control of the central bank to achieve its economic stabilization goals.

A&P estimate via maximum likelihood a macro-to-yield model, in the sense that macro factors a¤ect, but are not a¤ected by, monetary factors. Ang et al. (2005) improve that model by estimating a bidirectional model with one latent factor and two macro factors using MCMC. They report that the model forecasts better than the unrestricted VAR.

Rudebusch and Wu (2003) develop a no arbitrage macro-structural model with macro variables and latent monetary factors that jointly drive yields. They report that output shocks have a signi…cant impact on intermediate yields and curvature and that in‡ation surprises have large e¤ects on the level of the yield curve. They also …nd that including macro factors improve the forecasts of the usual latent factor models. Dai and Philippon (2004) estimate a no arbitrage VAR model with one latent factor and government de…cit, in‡ation and real activity. They argue that the de…cit is an important factor behind the yield curve. Nelson-Siegel models are discussed by Diebold et al. (2006).

Wu (2003) considers the impact of macro shocks on U.S. term structure using a structural VAR model, and concludes that monetary-policy and aggregate-supply shocks are important determinants of the slope and level of the yield curve, respectively. Hejazi (2000), on the other hand, examines whether information in the U.S. yield curve can be useful for predicting monthly industrial output. Using a GARCH-M model, he shows that while T-bill spreads contain little or no predictive content, increases in term premiums, which are linear functions of the conditional variance of excess returns, have predictive content.

Kalev and Inder (2006) and Chen (2001) test the rational expectations theory using U.S. term structure data. The former authors investigate how much information about the future yields is contained in the current spot rates, and their results suggest that a signi…cant amount of freely available information is not incorporated in forming agents' expectations. The latter author incorporates in‡ation in a model that allows for changes in regime, and concludes that the regime-switching model does not reconcile the data with the expectations hypothesis.

All these papers use discrete-time model and monthly or quarterly data.

To summarize, our objectives include: i) assessing the importance of macro variables in an affine term structure model for a Brazilian sample; ii) given the high number of parameters, evaluating the imposition of restrictions; iii) estimating the effect of the identified shocks, such as the exchange rate and inflation, on the yield curve, and vice-versa.

We find that macro factors improve the performance of the models. Also, variance decompositions show that the macro variables are important factors for yield curve movements in the Brazilian local market, a result that is similar to that of Diebold et al. (2006) for the U.S. bond market.

Finally, we remark that identification has two meanings here. First, it is used as in the financial context as the elimination of free parameters that cannot be estimated, as in Dai and Singleton (2000). Second, it is used as in the VAR literature as the exogeneity ordering of the state variables.

2. Affine Model

2.1. Continuous-Time. A probability space $(\Omega, \mathcal{F}, P)$ is fixed and no arbitrage is assumed. The price at time t of a zero-coupon bond paying \$1 at maturity date $t+\tau$ is $P(t,\tau) = E^{Q}\left[\exp\left(-\int_t^{t+\tau} r_t dt\right) \mid \mathcal{F}_t\right]$. The conditional expectation is taken under the equivalent martingale measure Q, and r_t is the stochastic instantaneous discount rate. Below, we discuss the pricing equations (see also Duffie, 2001).

The vector $X_t \in \mathbb{R}^p$ represents the state of the economy, and the short rate and risk premium process are given by time-varying processes $r_t = \delta_0 + \delta_1 \cdot X_t$ and $\lambda_t = \lambda_0 + \lambda_1 \cdot X_t$. It is assumed that X_t follows a Gaussian process with mean reversion. Under the objective P-measure,

$$dX_t = K(\xi - X_t)dt + \Sigma dw_t. \tag{2.1}$$

The $p \times p$ and $p \times 1$ parameters K and ξ represent the mean reversion coefficient and the long-term mean short rate, and $\Sigma\Sigma^T$ is the instantaneous variance-covariance matrix of the p-dimensional standard Brownian shocks w_t.

By Girsanov, under the martingale measure Q, $dX_t = K^{\star}(\xi^{\star} - X_t)dt + \Sigma dw_t^{\star}$, where $dw_t^{\star} = dw_t + \lambda_t dt$ is a standard Q-Brownian motion, and $K^{\star} = K + \Sigma\lambda_1$, $\xi^{\star} = K^{\star -1}(K\xi - \Sigma\lambda_0)$.

Multifactor Feynman-Kac formula states that, given technical conditions, if $E^{Q}\left[\exp\left(-\int_t^{t+\tau} r(X_u)du\right) \mid \mathcal{F}_t\right] = v(X_t,t,\tau)$, then $v(x,t,\tau)$ must satisfy $\mathcal{D}v(x,t,\tau) - r(x)v(x,t,\tau) = 0$, $v(x,t,0) = 1$, where the operator $\mathcal{D}$ is given by

$$\mathcal{D}v(x,t,\tau) := v_t(x,t,\tau) + v_x(x,t,\tau) \cdot K^{\star}(\xi^{\star} - x) + \frac{1}{2} tr[\Sigma\Sigma' v_{xx}(x,t,v)].$$

Applying Feynman-Kac to our pricing equation, it turns out that $P(t,\tau,X_t) = e^{\alpha(\tau)+\beta(\tau)\cdot X_t}$, where

$$\beta'(\tau) = -\delta_1 - K^{\star\prime}\beta(\tau), \tag{2.2}$$

$$\alpha'(\tau) = -\delta_0 + \xi^{\star\prime}K^{\star\prime}\beta(\tau) + \frac{1}{2}\beta(\tau)'\Sigma\Sigma'\beta(\tau). \tag{2.3}$$

are Riccati differential equations whose explicit solutions exist only in some special cases, such as when K is diagonal, but Runge-Kutta numerical integration efficiently solves equations (2.2) and (2.3).

The yield function is $Y(t,\tau) = -\frac{\alpha(\tau)}{\tau} - \frac{\beta(\tau)}{\tau} \cdot X_t$, or, defining $A(\tau) = -\frac{\alpha(\tau)}{\tau}$ and $B(\tau) = -\frac{\beta(\tau)}{\tau}$, $Y(t,\tau) = A(\tau) + B(\tau) \cdot X_t$. Stacking the equations for the n yield maturities, we arrive at a more concise expression:

$$Y_t = A + BX_t, \tag{2.4}$$

where $Y_t = (Y(t,\tau_1), ..., Y(t,\tau_n))'$.

Let T be the number of observations. The log-likelihood is the log of the density function of the sequence of observed yields $(Y_1, ..., Y_T)$. To calculate it we must first find the transition density of $X_{t_i|t_{i-1}}$, by integrating the equation (??):

$$X_{t_i|t_{i-1}} = (1 - e^{-K(t_i - t_{i-1})})X_{t_{i-1}} + e^{-K(t_i - t_{i-1})}\xi + \int_{t_{i-1}}^{t_i} e^{-K(t_i - u)}\Sigma\, dw_u. \tag{2.5}$$

The stochastic integral term above is Gaussian with zero mean and variance

$$E\left[\int_{t_{i-1}}^{t_i} e^{-K(t_i-u)}\Sigma\, dw_u\right]^2 = \int_{t_{i-1}}^{t_i} e^{-K(t_i-u)}\Sigma\Sigma'(e^{-K(t_i-u)})' du. \tag{2.6}$$

Hence, $X_{t_i|t_{i-1}} \sim N(\mu_i, \sigma_i^2)$, where $\mu_i = (1 - e^{-K(t_i - t_{i-1})})X_{t_{i-1}} + e^{-K(t_i - t_{i-1})}\xi$ and σ_i^2 is the above integral. When $dt = t_i - t_{i-1}$ is small, which is the case with daily data, the integral (2.6) can be well approximated using

$$\sigma_i^2 \simeq e^{-Kdt}\Sigma\Sigma'(e^{-Kdt})' dt, \tag{2.7}$$

and we have $X_{t_i|t_{i-1}} = \mu_i + \sigma_i N(0, I)$,with $\sigma_i = e^{-Kdt}\Sigma\sqrt{dt}$.

Now suppose the vectors X_t and Y_t have the same dimension, that is, the number of yield maturities equals the number of state variables. Then, we can invert the linear equation (2.4) and find X_t as a function h of Y_t: $X_t = B^{-1}(Y_t - A) = h(Y_t)$. Using change of variables, it follows that

$$\log f_Y(Y_{t_1}, ..., Y_{t_T}; \theta) = \sum_{i=2}^{T} (\log f_{X_{t_i}|X_{t_{i-1}}}(X_{t_i}; \theta) + \log|\det \nabla h|). \tag{2.8}$$

If we want to use additional yields, the direct inversion is not possible (a fact known as "stochastic singularity"). This problem is circumvented following Chen and Scott (1993), adding measurement errors to the extra yields.

Let Y_t^1 denote p out of n maturities to be priced without error. The other yields are denoted by Y_t^2, and they will have independent normal measurement errors $u(t,\tau) \sim N(0, \sigma^2(\tau))$. This is the method chosen for our continuous time versions. Thus,

$$Y(t,\tau) = A(\tau) + B(\tau) \cdot X_t + \sigma_u u_t.$$

The model depends on the set of parameters $\theta = (\delta_0, \delta_1, K, \xi, \lambda_0, \lambda_1, \Sigma, \sigma_u)$.

The Gaussian affine model has constant volatility and is the simplest specification of the affine family. It was chosen since the inclusion of macro factors substantially complicates the estimation of the parameters given the scarcity of data. Also, note that macro factors such as the VIX (the Chicago Board volatility index calculated from S&P 500 stock index option prices) can approximately take the role of stochastic volatility for Gaussian models.

2.1.1. Adding Macro Factors. Let now $X_t = (M_t, \theta_t)$, where M_t and θ_t denote vectors with p macro and q latent variables. The dynamics of the state vector is

$$\begin{bmatrix} dM_t \\ d\theta_t \end{bmatrix} = \begin{bmatrix} K_{MM} & K_{M\theta} \\ K_{\theta M} & K_{\theta\theta} \end{bmatrix} \left(\begin{bmatrix} \xi_M \\ \xi_\theta \end{bmatrix} - \begin{bmatrix} M_t \\ \theta_t \end{bmatrix} \right) + \begin{bmatrix} \Sigma_{MM} & 0 \\ \Sigma_{\theta M} & \Sigma_{\theta\theta} \end{bmatrix} \begin{bmatrix} dw_t^M \\ dw_t^\theta \end{bmatrix},$$

and the short rate equation combines a Taylor Rule and an affine model, $r_t = \delta_0 + \delta_{11}' M_t + \delta_{12}' \theta_t$. This permits the study of the inter-relations between macroeconomic questions, such as inflation target in monetary policy, and finance problems, such as derivative pricing, while affine tractability is retained. In fact, the pricing equations are simply higher dimensional versions of the earlier equations:

$$Y(t,\tau) = A(\tau) + B^M(\tau)' M_t + B^\theta(\tau)' \theta_t + \sigma_u u_t,$$

where A, B are still solutions of Riccati equations.

The likelihood is calculated as follows. As discussed before, Y_t^1 and Y_t^2 denote yield maturities without and with measurement errors u_t. We have

$$\begin{bmatrix} M_t \\ Y_t^1 \\ Y_t^2 \end{bmatrix} = \begin{bmatrix} 0 \\ A^1 \\ A^2 \end{bmatrix} + \begin{bmatrix} I & 0 & 0 \\ B^{M1} & B^{\theta 1} & 0 \\ B^{M2} & B^{\theta 2} & I \end{bmatrix} \begin{bmatrix} M_t \\ \theta_t \\ u_t \end{bmatrix}. \tag{2.9}$$

Denote by h the function that maps the state vector (M_t, θ_t, u_t) to (M_t, Y_t^1, Y_t^2). One obtains θ_t inverting on Y_t^1, $\theta_t = (B^{\theta 1})^{-1}(Y_t^1 - A^1 - B^{M1\prime} M_t) = h(Y_t^1, M_t)$, and u_t by solving for it in the last equation. Thus,

(2.10)

$$\log f_Y(Y_{t_1}, ..., Y_{t_T}; \Theta) = \log f_X(X_{t_1}, ..., X_{t_T}; \Theta) + \log f_u(u_{t_1,...,}u_{t_T}) + \log|\det \nabla h|^{T-1}$$

$$= (T-1)\log|\det B^{\theta 1}| + \sum_{t=2} \log f_{X_t|X_{t-1}}(X_t; \Theta) + \log f_u(u_t).$$

In a model in which the macro factors are not affected by the yield curve like Ang and Piazzesi (2003), the macro factors can be estimated separately in a first step.

2.2. Discrete-Time. Following the A&P approach, we derive the discrete time equations. Again, no arbitrage is assumed. The price at time t of a zero coupon bond maturing $s+1$ periods ahead is $p_t^{s+1} = E^{\mathbb{Q}}[\exp(-r_t) p_{t+1}^s | \mathcal{F}_t]$, where, as before, $\mathbb{Q} \sim \mathbb{P}$ is the martingale measure, $\mathcal{F}_t$ the filtration, and $r_t = \delta_0 + \delta_1 X_t$ and $\lambda_t = \lambda_0 + \lambda_1 X_t$ are the short rate and risk premium equations. The dynamics of the state vector is the multifactor autoregression $X_t = \mu + \Phi X_{t-1} + \Sigma \epsilon_t$.

Denote by ξ_t the Radon-Nikodym derivative $\frac{dQ}{dP} = \xi_t$. A discrete-time "version" of Girsanov theorem is assumed setting $\xi_{t+1} = \xi_t \exp(-\frac{1}{2}\lambda_t' \lambda_t - \lambda_t \epsilon_{t+1})$, where $\{\epsilon_t\}$ are independent normal errors. Then, the pricing kernel becomes $m_t = \exp(-r_t)\frac{\xi_{t+1}}{\xi_t}$, so that, by induction, one proves that the price of the bond is an exponential affine function of the state vector, i.e., $p_t^n = \exp(\alpha_n + \beta_n' X_t)$, where

$$\begin{aligned} \beta_{n+1}' &= -\delta_1'(1 + \Phi^\star + .. + \Phi^{\star n}), \\ \alpha_{n+1} &= -\delta_0 + \alpha_n + \beta_n' \mu^\star + \frac{1}{2}\beta_n' \Sigma\Sigma' \beta_n, \end{aligned} \tag{2.11}$$

with initial condition $\alpha_1 = -\delta_0$, $\beta_1 = -\delta_1$, and $\Phi^\star = \Phi - \Sigma\lambda_1$, $\mu^\star = \mu - \Sigma\lambda_0$. Then $Y_t^n = -\log p_t^n / n = A_n + B_n' X_t$, where $A_n = -\alpha_n/n$ and $B_n = -\beta_n/n$. Forming a vector of yields, we arrive at the same expression as in the continuous case, $Y_t = A + BX_t$.

The procedure to include macro factors is the same as in the continuous-time case. The yield curve is described by means of the state vector $X = (M, \theta)$. The observation equation relates the evolution of the yield curve to the state through matrices A and B, whose coefficients depend on the monetary rule that determines the short rate given the state of the economy, the affine risk premium, and the idiosyncratic variance error σ_1:

$$
\begin{aligned}
Y_t &= A(\delta_0, \Sigma\Sigma', \mu^\star, \Phi^\star) + B(\delta_1, \Phi^\star)X_t + \sigma u_t, \\
X_t &= \mu + \Phi X_{t-21} + \Sigma \epsilon_t, \\
r_t &= Y_t^1 = \delta_0 + \delta_1 X_t + \sigma_1 u_t^1,
\end{aligned}
\tag{2.12}
$$

where $\mu^\star = \mu - \Sigma'\lambda_0$, $\Phi^\star = \Phi - \lambda_1'\Sigma$. The parameters (μ, Φ) characterize the P-dynamics of the state variables, (δ_0, δ_1) the monetary rule that determines the short rate given the state of the economy, (λ_0, λ_1) the risk premiums describing the dynamics of the cross-section, and $\Sigma\Sigma'$ the covariance among the shocks. As in Johannes and Polson (2003), $(\mu^\star, \Phi^\star)$ are directly estimated, from which the premium is inferred.

In order to identify monetary factors, we imposed the condition (2.17) discussed below in each iteration, implying that only a subset of the elements of Σ is free. The parameters are $\Theta = (\mu, \Phi, \sigma, \theta, \zeta)$, where $\zeta = (\delta_0, \delta_1, \mu^\star, \Phi^\star, \Sigma\Sigma')$.

In the discrete-time case, normal measurement errors u_t are added to all maturities since Kalman filter is used.

Another difference is that we estimate a monthly model using daily data, by choosing a 21 days lag:

$$
X_t = \mu + \Phi X_{t-21} + \Sigma \epsilon_t, \ \epsilon_t \sim N(0, I). \tag{2.13}
$$

2.2.1. Impulse Response and Variance Decomposition. Impulse response functions (IRF) and variance decompositions (VD) are used to analyze the impact of macro shocks on yields and default probabilities. In discrete-time case, the IRF is $X_t = \Sigma\varepsilon_t + \Phi\Sigma\varepsilon_{t-1} + \Phi^2\Sigma\varepsilon_{t-2} + \Phi^3\Sigma\varepsilon_{t-3} + \ldots$ When $Y_t = A + BX_t$, clearly the response of the yield curve Y_t to the shocks becomes

$$
\begin{array}{ccccc}
B\Sigma\varepsilon_t & B\Phi\Sigma\varepsilon_t & B\Phi^2\Sigma\varepsilon_t & B\Phi^3\Sigma\varepsilon_t & \ldots \\
t+0 & t+1 & t+2 & t+3 & \ldots
\end{array}
$$

In continuous time, we have

$$
X_{t_i|t_{i-k}} = e^{-K(t_i - t_{i-k})} X_{i-k} + \sum_{l=0}^{k-1} \int_{t_{i-k+l}}^{t_{i-k+l+1}} e^{-K(t_i - u)} \Sigma \, dw_u .
$$

Using the approximation (2.7), it follows that the response of X_t to a shock ε_t in a interval of time of dt is

$$
\begin{array}{ccccc}
\Sigma\sqrt{dt}\varepsilon_t & e^{-Kdt}\Sigma\sqrt{dt}\varepsilon_t & e^{-2Kdt}\Sigma\sqrt{dt}\varepsilon_t & e^{-3Kdt}\Sigma\sqrt{dt}\varepsilon_t & \ldots \\
t+0 & t+1 & t+2 & t+3 & \ldots
\end{array}
\tag{2.14}
$$

Similarly, the response of the yield Y_t is given by

$$
\begin{array}{ccccc}
B\Sigma\sqrt{dt}\varepsilon_t & Be^{-Kdt}\Sigma\sqrt{dt}\varepsilon_t & Be^{-2Kdt}\Sigma\sqrt{dt}\varepsilon_t & Be^{-3Kdt}\Sigma\sqrt{dt}\varepsilon_t & \ldots \\
t+0 & t+1 & t+2 & t+3 & \ldots
\end{array}
\tag{2.15}
$$

To find the variance decomposition, note first that, in discrete time, the Mean Squared Error (MSE) of the s-periods ahead error $X_{t+s} - EX_{t+s|t}$ is calculated as follows:

$$MSE = \Sigma\Sigma' + \Phi\Sigma\Sigma'\Phi' + \Phi^2\Sigma\Sigma'(\Phi^2)' + ... + \Phi^s\Sigma\Sigma'(\Phi^s)'.$$

The contribution of the j-th factor to the MSE of X_{t+s} will be then be

$$\Sigma_j\Sigma_j' + \Phi\Sigma_j\Sigma_j'\Phi' + \Phi^2\Sigma_j\Sigma_j'(\Phi^2)' + ... + \Phi^s\Sigma_j\Sigma_j'(\Phi^s)'.$$

The j-th factor's contribution to the MSE of Y_{t+s} is

$$B\Sigma_j\Sigma_j'B' + B\Phi\Sigma_j\Sigma_j'\Phi'B' + B\Phi^2\Sigma_j\Sigma_j'(\Phi^2)'B' + ... + B\Phi^s\Sigma_j\Sigma_j'(\Phi^s)'B'.$$

In continuous time, it turns out that the s-period ahead MSE of is the integral:

$$MSE = \int_t^{t+s} e^{-K(t+s-u)}\Sigma\Sigma'(e^{-K(t+s-u)})'du.$$

Hence, the contribution corresponding to the j-th factor in the variance decomposition of X_{t+s} and Y_{t+s} at time t are

$$\int_t^{t+s} e^{-K(t+s-u)}\Sigma_j\Sigma_j'(e^{-K(t+s-u)})'du, \quad B'\left(\int_t^{t+s} e^{-K(t+s-u)}\Sigma_j\Sigma_j'(e^{-K(t+s-u)})'du\right)B. \tag{2.16}$$

2.3. **Model Specification.** The discrete (D) and continuous-time (C) versions differ in two aspects. In (C), the latent factor is defined using Chen-Scott inversion, in which we choose some yield maturities to be exactly priced, and the transition equation has a one-day lag.

The discrete version was specified admitting that all maturities have observation errors, and using the Kalman Filter algorithm to estimate the latent factor. The lag is chosen to be 21 days, which equals the average number of commercial days in a month. This lag smooths intra-monthly seasonalities. The estimator will not take into account the serial correlation that appears with the lag size, but the associated loss in efficiency disappears with longer series. Also, the correlation does not produce bias.

In summary, the model has two time dimensions, the historical time in which the sample is collected, and the time of the maturities of the yield curve. The first is always daily, while the latter was fixed as 1-day in (C) and as 1-month in (D). The difference of daily or monthly transition affects the definition of Φ.

In order to identify the model, we impose

$$\Sigma = \begin{pmatrix} \Sigma^{MM} & 0 \\ 0 & I \end{pmatrix}, \quad EX = \overline{X} = \begin{pmatrix} \overline{X}_M \\ 0 \end{pmatrix}, \quad \Phi = \begin{pmatrix} \Phi^{MM} & \Phi^{M\theta} \\ \Phi^{\theta M} & \Phi^{\theta\theta} \end{pmatrix}, \tag{2.17}$$

where $\Phi_{\theta\theta}$ is lower triangular. It is shown in chapter 3 that this specification is exactly identified.

3. Inference

3.1. **Continuous-time case.** In the (C) version, we find the parameters by maximizing the log-likelihood with respect to the parameters, given the data. This estimation method produces asymptotically consistent, non-biased and normally distributed estimators. Let $L = \log f_Y$ denote the log-likelihod. When $T \to \infty$, $\hat{\theta} \to \theta$ a.s. and $T^{\frac{1}{2}}(\hat{\theta} - \theta) \to N(0, V)$ in distribution, where $V^{-1} =$

$E\left(\frac{\partial L(Y;\Psi)}{\partial\Psi}\frac{\partial L(Y;\Psi)'}{\partial\Psi}\right)$, or $-E\left(\frac{\partial^2 L(Y;\Psi)}{\partial\Psi^2}\right)$ using the information inequality. An estimator for V^{-1} is the empirical Hessian

$$\hat{V}^{-1} := -\frac{1}{T}\sum_{t=1}\left(\frac{\partial^2 L_t(Y|\Psi)}{\partial\Psi^2}\right),$$

where L_t represents the likelihood of the vector with t elements. More details can be found in Davidson and Mackinnon (1993), chapter 8.

Confidence intervals for the parameter estimates are found using the empirical Hessian. If the number of observations T is large enough, then the variance of $\hat{\Psi}-\Psi$ will be given by the diagonal of $N(0, V/T)$. Alternatively, one could obtain confidence intervals via simulation.

Our estimation strategy consisted in many trial optimizations using Matlab. Beginning with more restricted models, different starting vectors are chosen in numerical optimization trials until stable results are obtained.

The resultant parameters are posteriorly used in the initial vectors of the higher dimensional models maximization. Independent new trial maximizations from random vectors are also conducted and compared to other results.

In the end, the maximal results are chosen. Although this procedure may be path-dependent, the "curse of dimensionality" does not allow the use of a complete grid of random starting points as would be desirable. However, (C) models results can be checked against (D) models results, which are estimated through an entirely different method.

3.2. Discrete-time case. Version (D) is estimated via MCMC, a Bayesian approach, which obtains the joint distribution $f(\Psi,\theta|M,Y)$ of the parameters and latent variables conditional on observed data.

The description of the MCMC involves i) the presentation of the Gibbs sampling and Metropolis-Hastings algorithms; ii) the Clifford-Hammersley theorem; and iii) Markov process limit theorems.

General references for this estimation method are Robert and Casella (2005), Gamerman and Lopes (2006), and for the specific case of financial econometrics, the survey by Johannes and Polson (2003).

Although $f(\Psi,\theta|M,Y)$ is generally unknown and extremely complex, the Clifford-Hammersley theorem guarantees that if the positivity condition is satisfied, it can be uniquely characterized by the lower dimensional complete conditional distributions $f(\Psi|M,Y,\theta)$ and $f(\theta|M,Y,\Psi)$, which, in turn, can be characterized by even lower dimensional complete distributions. For instance, if the set of parameters is divided into subsets, $\Psi=(\Psi_1,..,\Psi_n)$, the complete distributions $f(\Psi_i|\Psi_{-i},M,Y,\theta)$ determines $f(\Psi|M,Y,\theta)$.

The MCMC method provides algorithms through which the full conditional distribution is recovered from lower dimensional ones. They avoid high dimensional nonlinear optimizations.

The Gibbs sampling algorithm sequentially samples and updates the set of complete conditional distributions, and generates a Markov chain whose invariant measure is $f(\Psi,\theta|M,Y)$.

Ergodic and the central limit theorems can be applied to give conditions under which chains formed by the Gibbs sampling converge to desired distribution. The positivity condition, besides technical conditions, suffice.

When complete conditionals are unknown, the Metropolis-Hastings (M-H) method is used instead. In this case, the sampling comes from a candidate distribution, whose realizations are accepted or not with a probability given by the ratio between the current and the new realization of the likelihood.

In practice, it is easier to obtain convergence with Gibbs sampling. Thus, we must carefully break the set of parameters $\Psi = (\delta_0, \delta_1, \mu, \Phi, \Sigma\Sigma', \mu^\star, \Phi^\star, \sigma_u, \theta)$ into convenient subsets which can be analytically sampled.

Details of the specific implementation of the algorithms to our models are given next. Subproblems 1-3 below are cases of Gibbs sampling, corresponding to, respectively, the estimation of a VAR model, the estimation of the variance of independent time series, and the joint distribution of latent factors. Subproblem (4), relative to $\zeta = (\delta_0, \delta_1, \mu^\star, \Phi^\star, \Sigma\Sigma')$, does not have known closed expressions and is sampled via M-H, with a proposal obtained from a normal or Wishart distribution, centered on the value of the previous iteration, and with an arbitrarily fixed variance such that the acceptance rate remains in the interval [0.2, 0.5].

The algorithm consists of the following steps. Given an initial vector (Ψ^0, θ^0), repeat for $k = 1..N$,

(1) Draw $(\mu^k, \Phi^k) \sim p(\mu, \Phi | \sigma^{k-1}, \zeta^{k-1}, \theta^{k-1}, Y, M)$,
(2) Draw $\sigma^k \sim p(\sigma | \mu^k, \Phi^k, \zeta^{k-1}, \theta^{k-1}, Y, M)$,
(3) Draw $\theta^k \sim p(\theta | \mu^k, \Phi^k, \sigma^k, \zeta^{k-1}, Y, M)$,
(4) Draw $\zeta_i^k \sim p(\zeta_i | \mu^k, \Phi^k, \sigma^k, \theta^k, \zeta_{-i}^{k-1}, Y, M)$.

More specifically, for the step k, we have:

Subproblem 1:

$$f(\mu, \Phi | \sigma^{k-1}, \zeta^{k-1}, \theta^{k-1}, Y, M) \sim N((X'X)^{-1}X'X^\star, (X'X)^{-1} \otimes \Sigma), \tag{3.1}$$

where $X = (X_1, ..., X_{T-1})'$, $X^\star = (X_2, ..., X_T)'$, $X = (M, \theta)$.

Subproblem 2:

$$f(\sigma | \mu^k, \Phi^k, \zeta^{k-1}, \theta^{k-1}, Y, M) \sim IG(diag(U'U/T), T), \tag{3.2}$$

where $U = Y - A - BX$, and IG is the inverse gamma distribution.

Subproblem 3:

$$f(\theta | \mu^k, \Phi^k, \sigma^k, \zeta^{k-1}, Y, M). \tag{3.3}$$

This problem is solved via the FFBS algorithm defined in the next subsection.

Subproblem 4:

$$f(\zeta_i | \zeta_{-i}^{k-1}, \mu^k, \Phi^k, \sigma^k, \theta^k, Y, M). \tag{3.4}$$

Here the M-H is used. Except for the $\Sigma\Sigma'$ case, we use random walk Metropolis: draw a candidate $\zeta_i^k \sim \zeta_i^{k-1} + N(0, c)$, where c is a constant. If

(3.5)
$$L(\zeta_i^k | \zeta_{-i}^{k-1}, \mu^k, \Phi^k, \sigma^k, \theta^k, Y, M) - L(\zeta_i^{k-1} | \zeta_{-i}^{k-1}, \mu^k, \Phi^k, \sigma^k, \theta^k, Y, M) > \log(z),$$

where L is the loglikelihood, detailed in the next section, and $z \sim U(0, 1)$, then accept ζ_i^k, otherwise $\zeta_i^k = \zeta_i^{k-1}$. Calibrating c, the acceptance ratio is maintained in the [20%, 50%] range.

In the $\Sigma\Sigma'$ case, we use independent random walk, where the candidate distribution is the inverted Wishart distribution.

3.3. Kalman filter and FFBS algorithm. This subsection gives details of the Kalman filter and the forward-filtering backward sampling (FFBS) algorithms of the dynamic linear model (DLM) in which part of the state vector is observed (M). See also West and Harrison (1997). We have

$$Y_t = A + BX_t + \sigma_u u_t,\ u_t \sim N(0, I), \text{ diagonal } \sigma,$$
$$X_t = \mu + \Phi X_{t-h} + \Sigma \epsilon_t,\ \varepsilon_t \sim N(0, I),$$
$$X_t = [M_t; \theta_t],$$

where A and B are given by (2.11).

The algorithm works as follows:

$$\Theta = (\delta_0, \delta_1, \Phi, \mu, \lambda_0, \lambda_1, \Sigma);$$
$$D_t = \{\Theta, Y_1, ..., Y_t, M_1, ..., M_t\};$$
$$\theta_0 \sim N(m_0, C_0) \text{ is given};$$
$$\text{Prior of the state variables: } X_t|D_{t-h} \sim N(a_t, R_t);$$
$$E(X_t|D_{t-h}) = a_t = \mu + \Phi m_{t-h};$$
$$V(X_t|D_{t-h}) = R_t = \Phi C'_{t-h}\Phi' + V;$$
$$\text{Forecast of the yields: } (Y_t|D_{t-h}) \sim N(f_t, Q_t)$$

where

$$f_t = A + Ba_t, \quad Q_t = BR_tB' + \sigma'\sigma;$$
$$\text{Posterior of the state variables: } (X_t|D_t) \sim N(m_t, C_t);$$
$$E(X_t|D_t) = m_t = (M_t, m_t^\theta);$$
$$V(X_t|D_t) = C_t = \begin{pmatrix} 0 & 0 \\ 0 & c_t^\theta \end{pmatrix};$$
$$m_t^\theta = a_t^\theta + R_t^\theta B^{\theta\prime} Q_t^{-1}(Y_t - f_t);$$
$$c_t^\theta = R_t^\theta + R_t^\theta B^{\theta\prime} Q_t^{-1} B^\theta R_t^{\theta\prime}.$$

The marginal log-likelihood L of the yields is

$$L = \log f(Y|\Theta, N) = \log \prod_t f(Y_t|D_{t-h}) = \sum_t -\frac{1}{2}[\log|Q_t| + (Y_t - f_t)Q_t^{-1}(Y_t - f_t)'].$$

The step 4 of the MCMC requires a realization of

$$\theta_t^w \sim \theta_t|D_T, t = 1, .., T, \tag{3.6}$$

which is obtained by FFBS. In what follows, we present a modification of the FFBS algorithm where part of the state variables is observable.

Carter and Kohn (1994) proved that the sampling of (3.6) is obtained by the reverse recursive sampling of $\theta_t^w \sim \theta_t|D_T, \theta_{t+1}$:

$$\theta_T^w|D_T \sim N(m_T, C_T),$$
$$\theta_t^w \sim N(h_t, H_t),$$

where

$$h_t = m_t + G_t(\theta_{t+1} - a_{t+1}),$$
$$H_t = C_t - G_t R_{t+1} G_t',$$
$$B_t = C_t \Phi' R_{t+1}^{-1}.$$

In the case with observed variables, we have:

$$G_t = \begin{bmatrix} 0 & 0 \\ 0 & c_t^\theta \end{bmatrix} \begin{bmatrix} \Phi_{mm} & \Phi_{m\theta} \\ \Phi_{\theta m} & \Phi_{\theta\theta} \end{bmatrix} R_{t+1}^{-1} = \begin{bmatrix} 0 & 0 \\ G_t^{\theta m} & G_t^{\theta\theta} \end{bmatrix},$$

$$h_t = \begin{bmatrix} M_t \\ m_t^\theta \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ G_t^{\theta m} & G_t^{\theta\theta} \end{bmatrix} \begin{bmatrix} M_{t+1} - a_{t+1}^m \\ \theta_{t+1}^w - a_t^\theta \end{bmatrix} = \begin{bmatrix} M_t \\ h_t^\theta \end{bmatrix},$$

$$H_t = \begin{bmatrix} 0 & 0 \\ 0 & c_t^\theta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ G_t^{\theta m} & G_t^{\theta\theta} \end{bmatrix} \begin{bmatrix} R_t^{mm} & R_t^{m\theta} \\ R_t^{\theta m} & R_t^{\theta\theta} \end{bmatrix} \begin{bmatrix} 0 & G_t^{\theta m\prime} \\ 0 & G_t^{\theta\theta\prime} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & H_t^\theta \end{bmatrix},$$

where

$$h_t^\theta = G_t^{\theta m}(M_{t+1} - a_{t+1}^m) + G_t^{\theta\theta}(\theta_{t+1}^w - a_t^\theta),$$

$$H_t^\theta = G_t^{\theta m} R_t^{mm} G_t^{\theta m\prime} + 2G_t^{\theta\theta} R G_t^{\theta m\prime} + G_t^{\theta\theta} R_t^{\theta\theta} G_t^{\theta\theta\prime},$$

$$\theta_t^w \sim N(h_t^\theta, H_t^\theta) \text{ repeated for } t = T - 1, .., 2.$$

4. Results

The expected inflation and the yield curve were extracted from contracts traded on BM&F Bovespa. Our sample contains DIxPRE swap contracts with maturities $\{1, 2, 3, 6, 9, 12, 18, 24, 36\}$-months, which represent the term structure, and INPCxDI swaps, which provide the difference between the inflation rate measured by the consumer price index and the floating interest rate observed at the contracted maturity. The ratio between the earnings of the latter asset and of the corresponding DIxPRE swap was taken as a measure of the expected inflation for that maturity. However, this ratio contains a risk premium that was supposed constant and disregarded. We chose the 6-month ahead expected inflation. Our sample was determined by the availability of those contracts at the time we collected the data, and it goes from April 2002 to October 2005, totaling 870 days. Figures 1 and 2 illustrate the evolution of the term structure and of the macro factors. As remarked in the introduction, it is difficult to find historical series that span many economic cycles in the Brazilian economy, because of the profound changes of regime that have occurred until recently.

Following Litterman and Scheinkman (1991), we analyzed the yield curve in Brazil, which indicated that 99% of the variance of the nine yield maturities in our sample can be described by two principal components (90% and 9% for the first and second component). This motivated us to fix two unobserved monetary factors. Then, the main sources of nominal shocks in the economy, the log of the nominal exchange rate and the expected inflation rate, were added. The independent structural shocks associated with those variables were identified supposing that the innovation of the exchange rate is more exogenous than the innovation of expected inflation, and zero correlation between the latent and observed shocks. Thus, the model has four independent exogenous shocks affecting the short rate.

Summary statistics are given in Table A.

In the A&P model, the relation between the short rate and the state variables is interpreted as a reaction function of the MA to changes in the state of the economy (Taylor rule). In our model, specified for a daily sample, this relation will also reflect the market reaction to new information arriving continuously between the MA meetings to decide the benchmark rate, regarding future changes by the MA and macroeconomic values. Even though other market conditions influence the

short rate, the systematic reaction of the MA to shocks affecting inflation and the exchange rate is contained in the impulse response function of the identified shocks.

Table A. Summary statistics

		Central moments			Autocorrelations		
	mean	std dev	skew	kurt	lag 21	lag 42	lag 63
1m	0.1961	0.0361	0.7789	2.2146	0.9524	0.8640	0.7372
2m	0.1975	0.0366	0.7195	2.1556	0.9458	0.8632	0.7480
3m	0.1990	0.0370	0.6570	2.0985	0.9403	0.8636	0.7609
6m	0.2034	0.0398	0.5898	2.0248	0.9196	0.8425	0.7665
9m	0.2070	0.0440	0.6654	2.0919	0.9040	0.8208	0.7536
12m	0.2104	0.0486	0.7671	2.2406	0.8998	0.8137	0.7463
18m	0.2168	0.0574	0.9009	2.4791	0.8973	0.8017	0.7329
24m	0.2227	0.0645	0.9586	2.5737	0.8973	0.7959	0.7248
36m	0.2330	0.0754	0.9973	2.6376	0.8945	0.7899	0.7149
EX	1.0870	0.1100	0.3787	3.2437	0.7616	0.5721	0.3460
inflation	0.0748	0.0311	-0.2446	3.0408	0.7991	0.6703	0.5394

Data description. The yield data is composed of daily rates obtained from the DIxPre swaps, provided by BM&F Bovespa, which approximates the Brazilian Government Bonds zero-coupon constant maturity rates. The exchange rate series is provided by IPEADATA. The inflation data refers to a daily six-month ahead expected inflation series constructed using the DIxPre and INPCxDI swaps. The table contains yield sample means, standard deviations, skewness, kurtosis and autocorrelations. The sample period goes from April 2002 to October 2005.

In the A&P model, the relation between the short rate and the state variables is interpreted as a reaction function of the MA to changes in the state of the economy (Taylor rule). In our model, specified for a daily sample, this relation will also reflect the market reaction to new information arriving continuously between the MA meetings to decide the benchmark rate, regarding future changes by the MA and macroeconomic values. Even though other market conditions influence the short rate, the systematic reaction of the MA to shocks affecting inflation and the exchange rate is contained in the impulse response function of the identified shocks.

The identified monetary factors have different characteristics. The unobserved factor 1 is highly correlated to the difference between the long and the short rate, and so we call it slope. The unobserved factor 2 is highly correlated to the mean value of the rates, and is denoted by level. This is shown in Tables 1 and 2. In Rudebusch and Wu (2003), the estimated monetary factors show similar characteristics. Their model contains two unobserved monetary factors - slope and level - a MA reaction rule, and a transition equation derived from a "macro structural" model.

They interpreted the innovation that increases all the yield rates as a shock on the preferences of the MA with respect to the level of inflation, that is, as an alteration of the inflation target. The innovation of the factor 2 was interpreted as an alteration of the monetary policy determinants, which could be caused by credit crunches, price misalignments or increases of risk perception because of events like the terrorist attack in the U.S. in 2001. In other words, it is an innovation that is not linked to a movement of inflation, but rather to recurring financial market crises. An example in Brazil was the crisis preceding the presidential election in

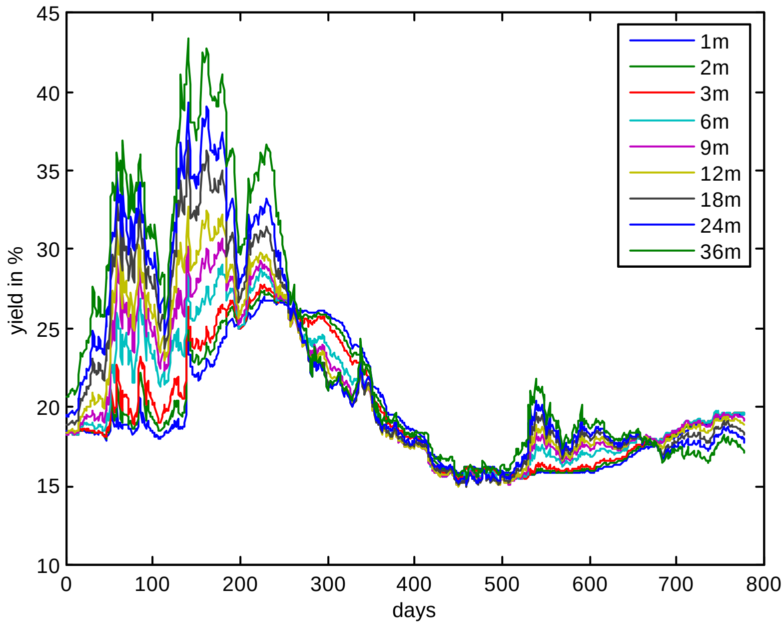


Figure 1. Evolution of the Brazilian domestic yield curve from April 2002 to October 2005.

2002, when local currency government bonds maturing after the election started to carry a spread due to the perceived risk of default.

We now comment about numerical issues in the estimation process. In the likelihood optimizations of (C) versions, lower dimensional models are estimated …rst and serve as initial points of other models. In the MCMC estimations of (D) versions, six chains are constructed and the one that presents the highest mean value of the log likelihood after convergence of the chain is chosen.

The analysis of the results includes three aspects: 1) the adherence, 2) the degree of the interdependence among macro shocks and yields, and 3) the dynamic e¤ect of the identi…ed shocks on the yield curve and on macro variables.

4.1. Evaluating Speci…cations.

The inclusion of the macro variables and the imposition of parameter restrictions will be considered, motivated by economic arguments or parsimony. Restrictions can be of two types: on the state variables transition equation or on the short rate equation.

In the bilateral model (B), the dynamics of the latent variables and observed variables are joined through the transition equation. The macro variables directly a¤ect the yield curve via the Taylor rule, or indirectly through the transition equation. In the yield-to-macro unilateral model (U), the macro variables do not a¤ect the transition of the latent variables ($©_{\theta M} = \lambda_{\theta M} = 0$), but a¤ect them through the Taylor rule. This is the speci…cation estimated by A&P.

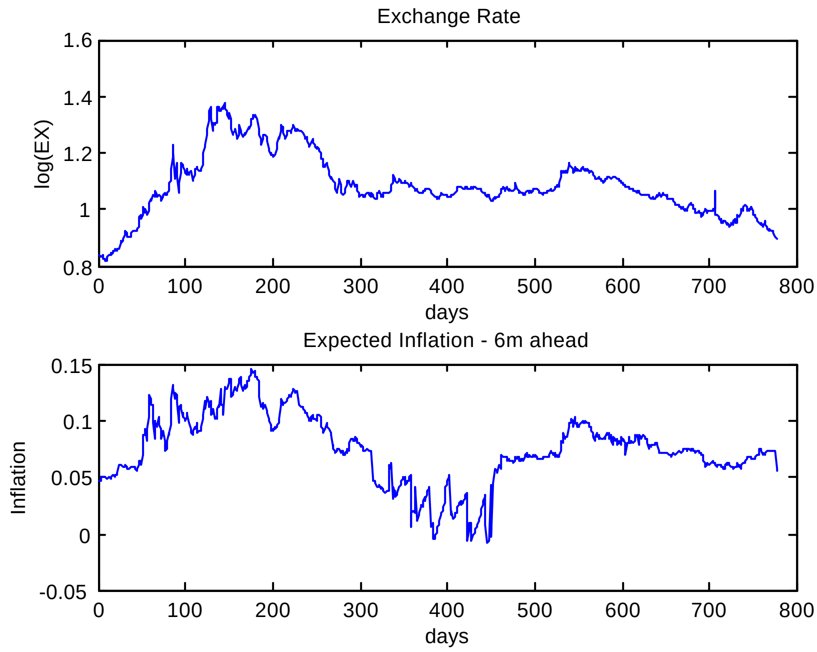


Figure 2. Historical series of the log of the exchange rate and of the 6-month ahead expected in‡ation as inferred from BM&F's swaps.

Ang et al. (2005) shows that restrictions in the short rate equation $r_t = \delta_0 + \delta_M M_t + \delta_\theta \theta_t$ emulate alternative monetary policies. We estimate three types of rule, the standard (s), the forward-looking (f) and the backward-looking (b):

- Standard Taylor rule: if $\delta_\theta = 0$, the MA reacts based on the present value of the variables.
- Forward-looking: if $\delta_M = 0$, the MA reacts based on the in…nite no discount future expectations of the macro variables.
- Backward-looking: if there are no restrictions, the MA smoothly reacts to past and present prices.

All the models have the same identi…cation assumption: contemporaneous macro and monetary shocks are not correlated, and the exchange rate is more exogenous than expected in‡ation.

Speci…cations with di¤erent number of macro variables are tested: (m) no macro, purely monetary: $X = \theta$; (i) including expected in‡ation: $X = (i, \theta)$; (e) including exchange rate: $X = (e, \theta)$; (ei) including both macro factors: $X = (e, i, \theta)$.

The maximizations of the likelihood in the (C) versions use numerical search algorithms, while in the (D) versions the mean value of the likelihood of the path is obtained after an initial number of iterations are discarded.

The models will are compared according to four criteria: 1) Akaike information = ¡2(log Likelihood/ # obser) + 2(# par/ # obser); 2) number of parameters; 3)

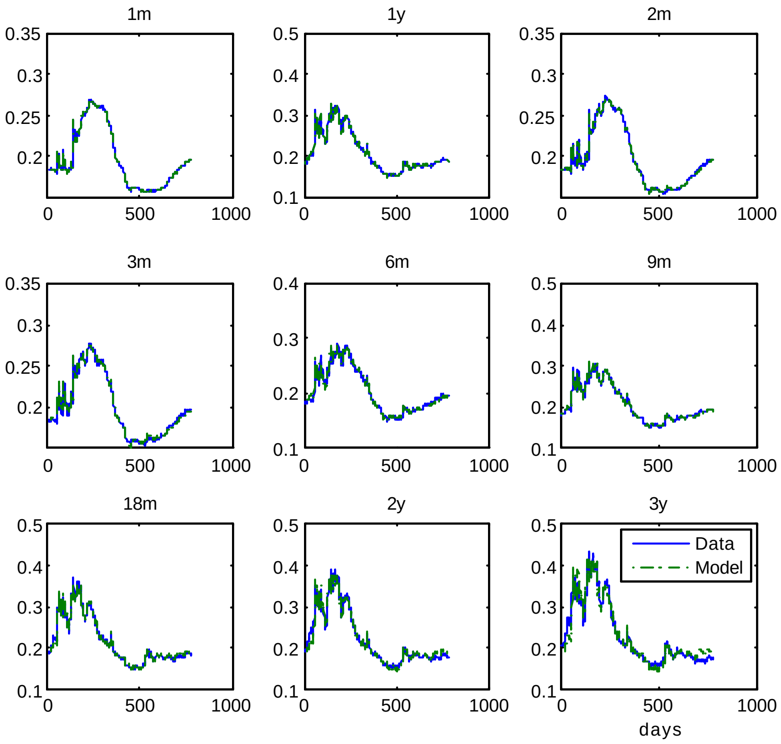


Figure 3. Compares the in-sample fitting of the continuous-time bilateral backward-looking specification with two macro and two latent factors, with actual data.

mean (of maturities) of normalized in-sample goodness-of-fit:

$$M(\ln) = \frac{1}{9}\sum_n \left(\frac{\sum_t (Y_t^n - \hat{Y}^n_{t|t-21})^2}{\sum_t (Y_t^n - Y^n_{t-21})^2} \right)^{\frac{1}{2}};$$

4) correlation between the factors and the level and slope of the yield curve; and 5) mean standard deviation of measurement error $M(\sigma)$. The Akaike information provides a normalized likelihood that penalizes the number of parameters.

We now discuss the results contained in Tables 1 and 2. The measurement errors and the in-sample adherence of both versions are generally similar. The monetary factor θ_1 is highly correlated to the slope and the monetary factor θ_2 to the level of the yield curve.

The Akaike information indicates that: i) the macro variables add information; ii) inflation is more informative than the exchange rate; iii) combined, the two macro variables are more informative; iv) in the discrete model with two macro factors, the bilateral model (B) is worse than the unilateral model (U); v) in the continuous case, the bilateral is only marginally higher than the unilateral; vi) the standard rule restriction is the worst speci…cation, and the backward-looking the best.

The models showed good …t in most of the speci…cations, as seen by the measurement errors. Figure 3 compares the model implied yield curves of the continuous-time backward-looking version with the data.

Table 1: Comparison of discrete-time speci…cations.

discrete	Akaike	#par	M(In)	$C(\theta_1$,slo)	$C(\theta_2$,lev)	$M(\sigma)$
yields only	-87	19	1.20	0.92	0.96	57
backward-looking Taylor rule: in‡ation						
unilateral	-108	28	0.97	0.97	0.74	63
bilateral	-105	32	0.96	0.87	1.00	42
backward-looking Taylor rule: exchange rate						
unilateral	-106	28	0.97	0.98	0.64	65
bilateral	-86	32	1.06	0.90	0.98	59
backward-looking Taylor rule: in‡ation, exchange rate						
unilateral	-111	42	0.92	0.99	0.94	128
bilateral	-104	50	0.91	0.94	0.86	56
standard Taylor rule: in‡ation, exchange rate						
bilateral	-84	48	1.35	0.90	0.86	175
forward-looking Taylor rule: in‡ation, exchange rate						
bilateral	-98	48	1.41	0.87	0.98	56

Summary of results of the discrete-time model. The …rst line corresponds to a purely monetary speci…cation and the others to speci…cations with in‡ation, exchange rate or both, with unilateral or bilateral dynamics. The Taylor rule can be backward, standard or forward-looking. The columns show the Akaike information, the number of parameters, the mean over the maturities of the in-sample model …tting normalized by the random walk …tting, M(In), the correlation between the factors and the slope, $C(\theta_1$,slo), or the level, $C(\theta_1$,lev), of the yield curve, and the mean measurement errors in basis points, $M(\sigma)$.

Table 2: Comparison of continuous-time speci…cations

discrete	Akaike	#par	M(In)	$C(\theta_1$,slo)	$C(\theta_2$,lev)	$M(\sigma)$
yields only	-90	17	0.96	0.90	0.68	45
backward-looking Taylor rule: in‡ation						
unilateral	-101	26	0.95	0.92	0.9	42
bilateral	-101	30	0.94	0.83	0.88	42
backward-looking Taylor rule: exchange rate						
unilateral	-99	26	0.97	0.93	0.85	44
bilateral	-99	30	0.95	0.70	0.86	43
backward-looking Taylor rule: in‡ation, exchange rate						
unilateral	-109	40	0.94	0.85	0.63	37
bilateral	-110	48	0.92	0.55	0.89	40
standard Taylor rule: in‡ation, exchange rate						
bilateral	-109	46	1.00	0.87	0.94	42
forward-looking Taylor rule: in‡ation, exchange rate						
bilateral	-110	46	0.93	0.50	0.91	40

Summary of results of the continuous-time model, containing the same items as Table1.

4.2. Comparing Model Dynamics. The (C) and (D) versions describe the interaction between the macro and the yield curve in distinct forms. As said before, in the continuous version some maturities are selected for the determination of the latent factors and the dynamics is daily, while the discrete version is a monthly model de…ned based on daily data, in which the latent factors are obtained through the Kalman …lter. We consider in this subsection six speci…cations that include the two macro factors to compare how the imposition of restrictions and choice of version alter the macro-yield interaction. All speci…cations were identi…ed considering that macro and latent shocks are contemporaneously uncorrelated, and that the exchange rate is more exogenous than expected in‡ation.

The variance decomposition of the speci…cations is contained in Tables 3 and 4. They show the proportion of the variance of the 18-month ahead forecast of the {1,9,36}-month yields and of the macro factors that are attributable to the monetary shocks and macro shocks, respectively.

In the unrestricted model, the macro variables a¤ect the yield curve directly through the Taylor rule, and indirectly through the state vector transition equation. The unilateral speci…cation eliminates the indirect channel, but the macro still affects the curve through the monetary policy channel. However, the decomposition of the yield movements of the backward-looking unilateral models showed no participation of the macro variables. This indicates that the macro-to-yield channel occurs mainly through the transition equation.

Table 3: Variance decomposition of macro shocks 18 months ahead

discrete	unilateral backward				bilateral backward				bilateral forward			
respnshock	ex	inf	slo	lev	ex	inf	slo	lev	ex	inf	slo	lev
ex	77	3	20	0	24	55	11	9	60	3	37	0
inf	3	58	16	23	3	81	14	3	2	32	63	3
1m	0	0	7	92	1	30	11	58	1	25	64	10
9m	0	0	23	77	2	45	16	37	0	10	85	5
3y	0	0	52	48	4	58	17	21	3	5	91	2

Variance decomposition of the exchange rate, in‡ation, and {1, 9, 36}-month yields using the discrete-time versions with two macro factors, backward or forward-looking Taylor rule and unilateral or bilateral dynamics. The lines contain the contributions of the exchange rate, in‡ation and latent factors.

Table 4: Variance decomposition of latent shocks 18 months ahead

discrete	unilateral backward				bilateral backward				bilateral forward			
respnshock	ex	inf	slo	lev	ex	inf	slo	lev	ex	inf	slo	lev
ex	31	05	55	10	77	04	19	00	78	3	19	00
inf	03	40	53	05	21	64	14	00	22	64	14	00
1m	00	00	78	22	22	24	20	34	21	21	21	36
9m	00	00	96	04	21	07	40	33	20	07	38	35
3y	00	00	00	00	28	01	47	24	27	02	45	26

Variance decomposition of the exchange rate, in‡ation and {1, 9, 36}-months yields using the continuous-time versions with two macro factors, backward or forward-looking Taylor rule and unilateral or bilateral dynamics. The lines contain the contributions of the exchange rate, in‡ation and latent factors.

Now, consider the second type of restriction, associated with the short rate formation. The forward- and backward-looking speci…cations show the same decomposition in the continuous case, while in the discrete case the e¤ect on macro variables is roughly the same.

For the bilateral models, the in‡ation shock a¤ects the variance of the yields, in accordance with the in‡ation target mechanism. Also, both versions show that the macro and latent variables are intertwined. Macro shocks impact the latent factors and vice-versa. However, we remark that in the continuous version, the impact in the macro-to-yield direction is greater, which is in accordance with Diebold et al. (2006).

4.3. Dynamic Properties of the Continuous-Time Unrestricted Model. During the sample period, the Brazilian MA successfully implemented an explicit in‡ation targeting regime. Hence, if our measure of market expectation of in‡ation is a good proxy for the measure of in‡ation used by the MA for policy decisions, then in‡ation would be an important factor to explain short rate movements. The other rates, being nominal variables, would also be a¤ected. Also, the in‡ation target remained relatively stable. With daily data, we interpret the determination of the short rate as the market updating the reaction function of MA, whose decisions only occur at certain intervals.

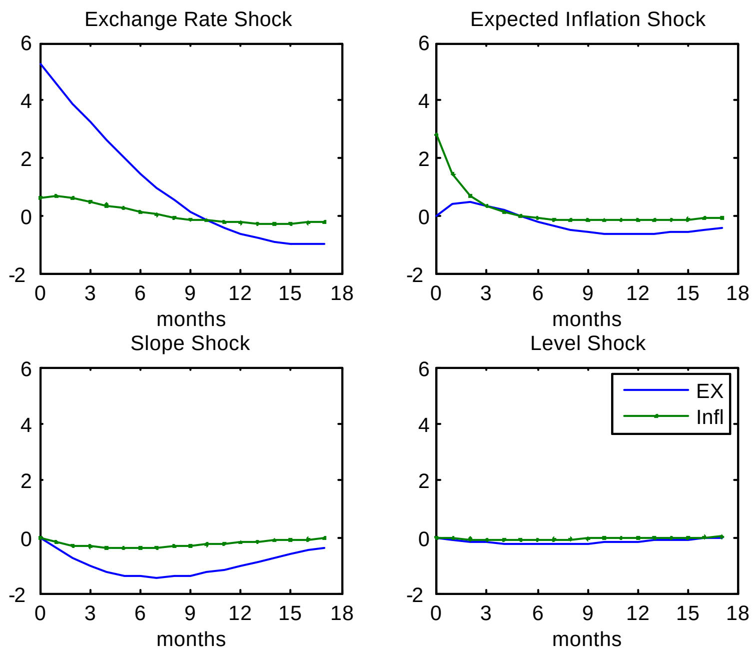


Figure 4. Impulse response functions of the continuous-time bilateral backward-looking speci…cation with two macro and two latent factors. Response of the macro factors.

Since a relatively short sample period is used, and since the frequency is daily, the model is more suited to describe the transmission mechanism from the exchange

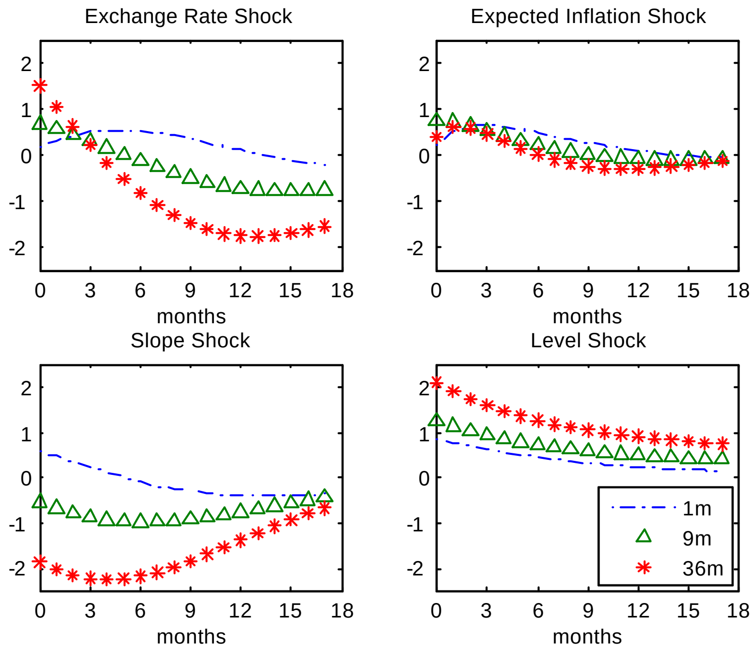


Figure 5. Impulse response functions of the continuous-time bilateral backward-looking speci…cation with two macro and two latent factors. Response of yields.

rate and in‡ation to interest rates (and vice-versa) in the short run. We chose the unrestricted continuous-time model to analyze details of this transmission, avoiding unnecessary restrictions and focusing on one type of model. The variance decomposition for the forecasting horizons of one and nine months are given in Table 5.

Table 5: Variance decomposition of the continuous-time version

	H=1m				H=9m			
respnshock	EX	in‡	slope	level	EX	in‡	slope	level
e	99	00	00	00	84	01	14	00
i	09	91	00	00	18	70	12	00
1m	05	14	24	57	24	29	09	38
9m	14	20	13	53	07	10	41	41
36m	17	03	39	41	11	02	58	30

Variance decomposition of the exchange rate, in‡ation and {1, 9, 36}-month yields. The model is the bilateral backward-looking continuous-time version with two macro factors one and nine month horizons. Macro and latent shocks.

The results reveal that exchange rate shocks are important for the Brazilian economy, corresponding to roughly 20% of the variation of in‡ation and interest rates. The in‡ation shock corresponds to another roughly 20% of the variation of the interest rates. Thus, macro factors are responsible for roughly 40% of interest rate movements, a proportion that is lower than the A&P macro-to-yield model and higher than the bilateral model of Diebold et al. (2006), both for the U.S. market. Ang et al. (2005) report a greater in‡uence of macro shocks on the spread of the rates (long rate minus short rate), but they use only one latent variable. Our estimation indicates that slope e¤ects, absent in single latent factor models, are very important.

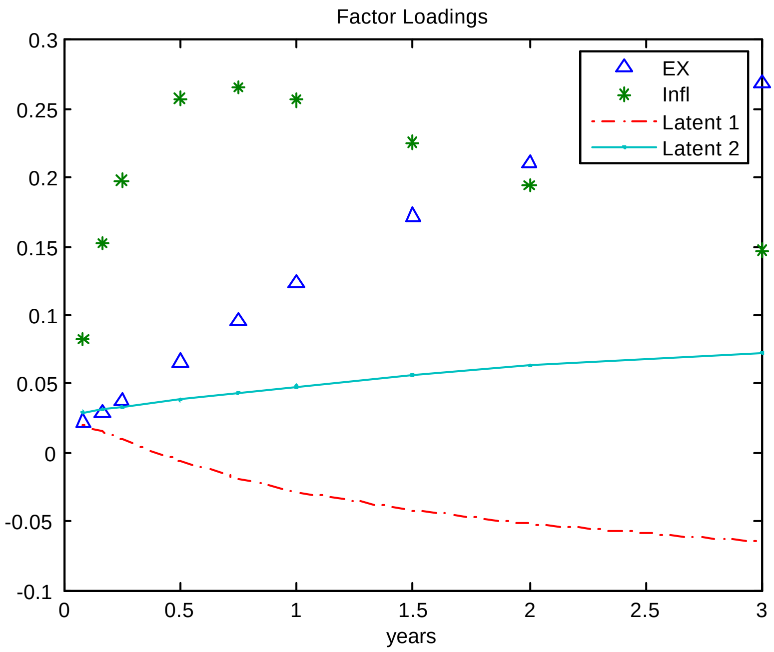


Figure 6. Shows the factor loadings of the continuous-time bilateral backward-looking speci…cation.

The market receives the expected in‡ation and exchange rate variations in different ways. The exchange rate has greater e¤ect on the long rates than in‡ation, indicating that the e¤ect of in‡ation on the short rate is perceived as more transitory. This represents a change in Brazilian market behavior in relation to the relatively recent past.

On the other hand, monetary shocks do not explain much of the macro variables, accounting for 12% over a nine-month horizon. Thus, the macro-to-yield direction is stronger, in accordance with the result obtained by Diebold et al. (2006) for the U.S. market.

The latent factors are still responsible for most of the yield variation, similar to the …ndings of Rudebusch and Wu (2003) and Ang et al. (2005). This is not to say macro variables are not important, since latent variables contain macro information by construction. Besides, they may also re‡ect omitted macro factors.

The impulse response functions are shown in Figures 4 and 5. Each graph contains the response of the short, medium and long rates to the indicated shock. The in‡ation shock is less expressive than other shocks; it equally raises all the rates and causes an initially upward and then downward impact on the exchange rate. The MA controls the short rate, but the long rate is formed in the market, suggesting that the market perceives that the rise is transitory.

The level shock raises all the nominal variables without changing the slope, as expected, followed by a smooth decrease over longer horizons. Indeed, if the in‡ation target regime works as an expectations organizer, than we would expect target changes to produce smooth rate changes.

The response to exchange rate shocks deeply alters the slope of the curve, raising the three-year yield, causing a strong but temporary e¤ect. The e¤ect on short rate is more modest but more persistent, which reveals the type of response of the MA.

Table 6: Parameters of the continuous-time bilateral backward-looking model with two macro factors

δ_1	0.019	-0.024	0.021	0.028			
	(0.007)	(0.016)	(0.002)	(0.001)			
$K^{\star}\xi^{\star}$	3.07	0.84	6.38	-6.22			
	(0.61)	(0.10)	(1.26)	(1.21)			
$K^{\star}$	1.12	27.54	3.38	0.29			
	(4.78)	(11.90)	(0.22)	(0.08)			
	-0.50	8.34	0.91	-0.04			
	(0.97)	(1.86)	(0.05)	(0.01)			
	1.02	-2.16	6.58				
	(6.44)	(13.90)	(0.50)				
	-4.00	-120.17	-7.89	-1.30			
	(14.33)	(27.47)	(0.70)	(0.25)			
K	1.99	-3.02	0.18	0.03			
	(1.27)	(4.57)	(0.11)	(0.11)			
	-1.22	7.99	0.10	0.02			
	(0.69)	(2.51)	(0.06)	(0.06)			
	-9.79	-34.46	0.41				
	(8.62)	(39.83)	(0.60)				
	3.93	-58.23	0.55	0.96			
	(6.60)	(29.32)	(0.58)	(0.60)			
§	0.18	0	0	0			
	(0.005)						
	0.02	0.10	0	0			
	(0.004)	(0.003)					
	0	0	1	0			
	0	0	0	1			
σ (in b.p.)	17.5	26.6	33.3	22.1	40.8	77.2	138
	(0.5)	(0.7)	(0.9)	(0.6)	(1.1)	(2.0)	(3.6)
$\delta_0 = \overline{Y}_1$							
$\xi = 0$							

Mean values, and the corresponding standard deviations in parenthesis, of the Monte Carlo parameter chains of the continuous-time bilateral backward-looking model with two macro factors.

Finally, the e¤ects of the slope shock also strongly alter the slope of the yield curve, being responsible for most of the variation of the yields. Rudebusch and Wu (2005) interpret this shock as a response to shocks not contained in the chosen macro variables.

Factor loading - matrix B - represents the e¤ect of state variables along the maturities and is presented in Figure 6. The level factor loading, as expected, is ‡at, and the slope factor loading decays over the maturities.

The parameter estimates of this section's speci…cation and their standard errors are given in Table 6.

5. Conclusion

This text follows the tradition of the …nance literature of using high-frequency data to estimate a¢ne models with macro variables. Methodological questions about inference, choice of speci…cation of latent factors and ad hoc parameter restrictions occurring in existing models motivated the consideration of a number of versions of the model. One group of versions is speci…ed in continuous-time with daily data, where latent factors are de…ned via Chen-Scott inversion, and estimated using maximum likelihood. Another group is speci…ed in discrete time with monthly dynamics, with factors extracted through the Kalman …lter, and estimated using MCMC. We estimated three of the monetary policy speci…cations proposed by Ang et al. (2005).

The models were used to analyze the Brazilian yield curve, which, because the country is an emerging market, has singular characteristics. The Brazilian economy evolved from a regime of high in‡ation and indexation up to 1994 to one of low in‡ation under a new currency, with a …xed exchange rate regime (a "sliding peg") that limited monetary policy options up to 1999, and then to an in‡ation targeting regime with a ‡oating exchange rate. This poses great challenges for the inference of the a¢ne models.

The main results are the following:

(1) The exchange rate and the expected in‡ation improve the model's capacity to explain yield curve movements.
(2) In spite of the cited di¤erences, the continuous and the discrete versions show qualitatively similar results in most cases.
(3) In general, restrictions on the number of parameters had a low e¤ect on the adherence, but signi…cantly altered the dynamic properties of the model. Care must be taken in the use of arbitrary restrictions.
(4) An important part of the variance of the yields is due to monetary factor shocks, which do not allow a direct interpretation, but is related to the level and slope of the yield curve and may accommodate omitted macro variables.
(5) The impulse response analysis showed that in‡ation shocks produce a temporary rise of moderate magnitude in the level, and exchange rate shocks produce signi…cant changes in the slope of the curve, mostly through the long-term yields.

CHAPTER 2

The Role of Macroeconomic Variables in Determining Sovereign Risk

1. Introduction

Sovereign risk is a subtype of credit risk related to the possibility of a government failing to honor its payment obligations. It is a fundamental component of emerging countries' yield curves. Sovereign risk is also very important for emerging market ...rms, since the cost of foreign ...nancing typically rises with the country risk. Accordingly, the following questions are of particular interest: What are the factors most a¤ecting the sovereign yield curve? Which variables have greatest impact on default probabilities? This study presents an empirical investigation of these questions by using an a¢ne term structure model with macroeconomic variables and default risk[1].

There are two main approaches in credit risk modeling: structural and reduced form models[2]. While the former provides a link between the probability of default and ...rms' fundamental variables, the latter relies on the market as the only source of information regarding ...rms' credit risk structure. Black and Scholes (1973) and Merton (1974) proposed the initial ideas concerning structural models based on options theory. Black and Cox (1976) introduced the basic structural framework in which default occurs the ...rst time the value of the ...rm's assets crosses a given default barrier. More recently, Leland (1994) extended the Black and Cox (1976) model, providing a signi...cant contribution to the capital structure theory. In his model, the ...rm's incentive structure determines the default barrier endogenously. That is, default is determined endogenously as the result of an optimal decision policy carried out by equity holders.

All the papers cited above deal with the corporate credit risk case. However, the sovereign credit risk di¤ers markedly from corporate risk[3]. For instance, it is not obvious how to model the incentive structure of a government and its optimal default decision, or what "assets" could be seized upon default. Moreover, post-default negotiating rounds regarding the recovery rate can be very complex and uncertain. Consequently, the use of structural models to assess the default risk of a country is a delicate question. Not surprisingly, it is di¢cult to ...nd studies of

[1]In this article, the term "macroeconomic (macro) variable" refers to any observable factor.

[2]Giesecke (2004) provides a short introductory survey about credit risk models.

[3]As discussed by Du¢e et al. (2003), the main di¤erences are: (i) A sovereign debt investor may not have recourse to a bankruptcy code at the default event. (ii) Sovereign default can be a political decision. (iii) The same bond can be renegotiated many times. (iv) It may be di¢cult to collateralize debt with assets into the country. (v) The government can opt for defaulting on internal or external debt. (vi) In the case of sovereign risk, it is necessary to take into account the role played by key variables such as exchange rates, ...scal dynamics, reserves in strong currency, level of exports and imports, gross domestic product, and in‡ation.

sovereign debt pricing based on the structural approach[4]. Therefore, we opt to use reduced models, where the default time is a totally inaccessible stopping time that is triggered by the …rst jump of a given exogenous intensity process[5]. This means that the default always comes as a "sudden surprise", which provides more realism to the model. In contrast, within the class of structural models, the evolution of assets usually follows a Brownian di¤usion, in which there are no such surprises and the default time is a predictable stopping time.

Lando (1998), and Du¢ e and Singleton (1999) develop versions of reduced models in which the default risk appears as an additional instantaneous spread in the pricing equation. The spread can be modeled using state factors. In particular, it can be incorporated into the a¢ ne framework of Du¢ e and Kan (1996), a widely used model o¤ering a good compromise between ‡exibility and numerical tractability[6]. Du¢ e et al. (2003) extend the reduced model to include the possibility of multiple defaults (or multiple "credit events", such as restructuring, renegotiation or regime switches). The model is estimated in two steps. First, the risk-free reference curve is estimated. Next, the defaultable sovereign curve is obtained conditional on the …rst stage estimates. As an illustration, they apply their model to analyze the term structure of credit spreads for bonds issued by the Russian Ministry of Finance (MinFin) over a sample period encompassing the default on domestic Russian GKO bonds in August, 1998. They investigate the determinants of the spreads, the degree of integration between di¤erent Russian bonds and the correlation between the spreads and the macroeconomic variables. Another paper applying reduced model to emerging markets is Pan and Singleton (2008), who analyze the sovereign term structures of Mexico, Turkey, and Korea through a dynamic approach.

Nevertheless, Du¢ e et al. (2003) and Pan and Singleton (2008) use a pure latent variables model. Thereby, the impact of macro factors changes on bond yields can be evaluated only indirectly through, for instance, a regression between observable and unobservable variables. Moreover, in pure latent models, the unobservable factors are abstractions that can, at best, be interpreted as geometric factors summarizing the yield curve movements, as shown by Litterman and Scheinkman (1991).

The modern literature linking the dynamics of the term structure with macro factors starts with Ang and Piazzesi (2003), who propose an ingenious solution to incorporate observable factors in the original framework of a¢ ne models. In their model, the macroeconomic factors a¤ect the entire yield curve. However, the interest rates do not a¤ect the macroeconomic factors, which means that monetary policy is ine¤ective. Similarly to Du¢ e et al. (2003), they employ a two-step estimation procedure, …rst determining the macro dynamics and then the latent dynamics conditional on the macro factors. Ang et al. (2007) also combine macro factors and no-arbitrage restrictions. Nevertheless, they use a Markov Chain Monte Carlo (MCMC) technique, which allows a single step estimation. On the other hand, Amato and Luisi (2006) estimate defaultable term structure models of corporate

---

[4] Exceptions are Xu and Ghezzi (2002) and Moreira and Rocha (2004).

[5] A stopping time is totally inaccessible if it can never be announced by an increasing sequence of predictable stopping times (see Schönbucher, 2003).

[6] An a¢ ne model is a multifactor dynamic term structure model, such that the state process X is an a¢ ne di¤usion, and the short short-term rate is also a¢ ne in X

bonds with the inclusion of macroeconomic variables following a conditional three-step procedure: macro factors, then U.S. yield curve, and in the end corporate bonds.

Following the advances brought by these previous studies, we examine the impact of macro factors on a defaultable term structure through an a¢ne model similar to that of Ang and Piazzesi (2003). We provide a comparison among a variety of speci…cations in order to determine the macro factors that most a¤ect credit spreads and default probabilities of an emerging country. We also use impulse response and variance decomposition techniques to analyze the direct in‡uence of observable macro factors on yields and default probabilities.

However, before estimating the parameters, one must choose an identi…cation strategy. Not all parameters of the multifactor a¢ne model can be estimated, since there are transformations of the parameter space preserving the likelihood. When sub-identi…ed, parameters can be arbitrarily rotated, while over-identi…ed speci…cations may distort the true response of the state variables. Based on the …ndings of Dai and Singleton (2000), we propose an identi…cation strategy for a¢ne models with macro factors and default.

We select Brazil as the case study. The reason for this choice is that Brazil is one of the most important emerging countries with a rich history of credit events[7]. When using Brazilian data, one must take into account that frequent regime switches have occurred until recently, such as change from very high in‡ation to a stable economy (Real Plan, July 1994), change from …xed to ‡oating exchange rate in a currency crisis in January 1999, and change of monetary policy to in‡ation targeting in July 1999. Thus, our sample comprises …ve and a half years of historical series. This sample size is compatible with that found in other recent academic studies containing data from emerging economies (see, for instance, Pan and Singleton, 2008, and Almeida and Vicente, 2009). Furthermore, following these authors, we decided to employ continuous-time modeling with high-frequency data in order to avoid small-sample biases.

Our main model contains …ve state variables: one latent factor for the reference default-free curve, one external macro factor, one internal macro factor, and two latent factors for the Brazilian sovereign yield curve. We test the following observable variables: Fed interest rates, VIX (index of implied volatility of options in the Standard & Poor's index), Brazilian Real/US Dollar exchange rates, São Paulo Stock Exchange index (IBovespa), and Brazilian interest rate swaps. In the estimation stage we follow common practice and use a two-step procedure as implemented by Du¢e et al. (2003).

In a nutshell, we contribute to the international empirical …nance literature in at least two aspects. First, we extend the works of Du¢e et al. (2003) and Pan and Singleton (2008) by incorporating macro variables in a dynamic term structure model with default risk. Second, our model allows a full interaction between latent and observable sovereign factors, which in a sense extends the study of Ang and Piazzesi (2003)[8].

---

[7] Jointly with India, Russia and China, Brazil is considered as among the faster growing developing economies in the world. Goldman Sachs refers to these countries as BRICs, an acronym that means Brazil, Russia, India and China (see Goldman Sachs, 2007).

[8] Diebold et al. (2006), using a simple statistical model, …nd strong evidence of two-way interaction between latent and macro factors.

Our main findings can be summarized as follows. First, VIX and Fed rates strongly affect the default probabilities in the short and in the long term, respectively. Second, VIX has a great effect on Brazilian sovereign yields, more than any investigated domestic macro indicator. This result agrees with one of Pan and Singleton's (2008) conclusions that VIX has the most explanatory power for Mexican credit default swap (CDS) spreads. Third, among the observable domestic factors only the slope of yield curve presents significant explanatory power of the Brazilian credit risk spread. Finally, a latent factor highly correlated with the level of the Brazilian sovereign curve predicts a substantial fraction of the yield and default probability movements. Since the Fed short rate has greater impact on the default probabilities than Brazilian domestic short rate, our model suggests that U.S. monetary policy is more important to the Brazilian term structure of credit spreads than the Brazilian monetary policy. We also assert that volatility of international market (measured in our model by VIX) is more important to determine Brazilian spread than local conditions. On the other hand, the moderate significance of the domestic yield curve slope indicates that expectations of Brazilian investors play an important role in determining the sovereign yield and default probabilities.

The rest of this article is organized as follows. In Section 2 we present the model. Section 3 describes the dataset used. Section 4 details the estimation procedure. Section 5 presents the results of implementing the dynamic models. Section 6 offers concluding remarks.

2. Affine Model with Default Risk and Macro Factors

Uncertainty in the economy is characterized by a filtered probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ where $(\mathcal{F}_t)_{t \geq 0}$ is a filtration generated by a standard N-dimensional Brownian motion $W^{\mathbb{P}} = \left(W_1^{\mathbb{P}}, \ldots, W_N^{\mathbb{P}}\right)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ (see Duffie, 2001). We assume the existence of a pricing measure $\mathbb{Q}$ under which discounted security prices are martingales with respect to $(\mathcal{F}_t)_{t \geq 0}$. The price P^D of a defaultable bond at time t that pays \$1 at maturity time $t + T$ is given by

$$P^D(t,T) = E_t^{\mathbb{Q}}\left[\mathbf{1}_{[\tau_d > t+T]} e^{\left(-\int_t^{t+T} r_u du\right)} + Z_{\tau_d} \mathbf{1}_{[\tau_d \leq t+T]} e^{\left(-\int_t^{t+T} r_u du\right)}\right], \tag{2.1}$$

where $\mathbf{1}_A$ is the indicator function of the set A. The first part of the right-hand side of (2.1) represents what the bondholder receives if the maturity time comes before the default time τ_d, a totally inaccessible stopping time. In case of default, the investor receives the random variable Z_{τ_d} at the default time. Lando (1998), and Duffie and Singleton (1999) prove that if τ_d is doubly stochastic with intensity η_t, the recovery upon default is given by $Z_{\tau_d} = (1 - \ell_{\tau_d})P^D(\tau_d, T)$, where ℓ_t is the loss rate in the market value, and if other technical conditions are satisfied, then

$$P^D(t,T) = E_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{t+T} (r_u + s_u)du\right)\right], \tag{2.2}$$

where $s_t = \ell_t \eta_t$ is the spread due to the possibility of default.

We now briefly explain the concept of doubly stochastic stopping time (for more details, see Schönbucher, 2003 or Duffie, 2001). Define $N(t) = \mathbf{1}_{[\tau_d \leq t]}$ as the associated counting process. It can be shown that $N(t)$ is a submartingale. Applying the Doob-Meyer theorem (see Shiryaev, 1995), we know there exists a predictable, non-decreasing process $C(t)$ called the compensator of $N(t)$. One property of the compensator is to give information about the probabilities of the jump time. The expected marginal increments of the compensator $dC(t)$ are equal to the probability of the default occurring in the next increment of time: $E_t^Q [C(t + \Delta t) - C(t)] = Q[N(t + \Delta t) - N(t) = 1|F_t]$. An intensity process η_t for $N(t)$ exists if it is progressively measurable and non-negative, and $C(t) = \int_0^t \eta(u)du$. Under regularity conditions, it turns out that

$$\eta(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} Q[\tau_d \leq t + \Delta t | \tau_d > t]. \tag{2.3}$$

Thus, $\eta(t)$ represents the evolution of the instantaneous probability of defaulting by $t + dt$ if default has not occurred up to t. Finally, τ_d is said to be doubly stochastic with intensity η if $N(t_2) - N(t_1)|\eta \sim \text{Poisson}\left(\int_{t_1}^{t_2} \eta(u)du\right)$. Therefore, in the reduced model, the default event is essentially given by the first jump of a Poisson process with stochastic intensity.

Our model is within the class of affine models analyzed by Duffie and Kan (1996). The state vector $X_t \in \mathbb{R}^N$ incorporates information about the United States, $X_t^{US} = \left(\theta_t^{US}, M_t^{US}\right)$, and Brazil, $X_t^{BR} = \left(M_t^{BR}, \theta_t^{BR}\right)$, that is, $X_t = \left(\theta_t^{US}, M_t^{US}, M_t^{BR}, \theta_t^{BR}\right)$, where the variables $\theta_t = \left(\theta_t^{US}, \theta_t^{BR}\right)$ and $M_t = \left(M_t^{US}, M_t^{BR}\right)$ represent latent and observable factors, respectively. In the affine model with default, s_t and r_t are specified as affine functions of the state vector. In other words, we assume that $s_t = \delta_0^s + \delta_1^s \cdot X_t$ and $r_t = \delta_0^r + (\delta_1^{r,US}, \delta_1^{r,BR}) \cdot X_t = \delta_0^r + \delta_1^r \cdot X_t$, where δ_0^s, $\delta_0^r \in \mathbb{R}$ and δ_1^s, $\delta_1^r \in \mathbb{R}^N$. Then the default-adjusted short-rate process is

$$R_t = R(X_t) = r_t + s_t = \delta_0^r + \delta_0^s + (\delta_1^r + \delta_1^s) \cdot X_t = \delta_0 + \delta_1 \cdot X_t. \tag{2.4}$$

The dynamics of the state variables is given by:

$$dX_t = \begin{bmatrix} d\theta_t^{US} \\ dM_t^{US} \\ dM_t^{BR} \\ d\theta_t^{BR} \end{bmatrix} =$$

$$\begin{bmatrix} K_{\theta,\theta}^{US,US} & 0 & 0 & 0 \\ K_{M,\theta}^{US,US} & K_{M,M}^{US,US} & 0 & 0 \\ K_{M,\theta}^{BR,US} & K_{M,M}^{BR,US} & K_{M,M}^{BR,BR} & K_{M,\theta}^{BR,BR} \\ K_{\theta,\theta}^{BR,US} & K_{\theta,M}^{BR,US} & K_{\theta,M}^{BR,BR} & K_{\theta,\theta}^{BR,BR} \end{bmatrix} \left(\begin{bmatrix} \xi_\theta^{US} \\ \xi_M^{US} \\ \xi_M^{BR} \\ \xi_\theta^{BR} \end{bmatrix} - \begin{bmatrix} \theta_t^{US} \\ M_t^{US} \\ M_t^{BR} \\ \theta_t^{BR} \end{bmatrix} \right) dt$$

$$+ \begin{bmatrix} \Sigma_{\theta,\theta}^{US,US} & 0 & 0 & 0 \\ \Sigma_{M,\theta}^{US,US} & \Sigma_{M,M}^{US,US} & 0 & 0 \\ \Sigma_{M,\theta}^{BR,US} & \Sigma_{M,M}^{BR,US} & \Sigma_{M,M}^{BR,BR} & 0 \\ \Sigma_{\theta,\theta}^{BR,US} & \Sigma_{\theta,M}^{BR,US} & \Sigma_{\theta,M}^{BR,BR} & \Sigma_{\theta,\theta}^{BR,BR} \end{bmatrix} \begin{bmatrix} dW_\theta^{P,US}(t) \\ dW_M^{P,US}(t) \\ dW_M^{P,BR}(t) \\ dW_\theta^{P,BR}(t) \end{bmatrix}$$

$$= K(\xi - X_t)dt + \Sigma dW^P(t), \tag{2.5}$$

where K and Σ are $N \times N$ matrices and $\xi \in \mathbb{R}^N$. That is, X follows an affine process with constant volatility. Similar to Duffie et al. (2003), we set a "block-triangular" form for the dynamics of the state variables. The zeros above the main diagonal of Σ and K imply that the American yield curve factors affect the Brazilian yield curve factors, but not vice versa. Furthermore, unlike Ang and Piazzesi (2003), we allow the macro and yield factors to interact fully.

The connection between martingale probability measure Q and objective probability measure P is given by Girsanov's Theorem with a time-varying risk premium:

$$dW_t^{P} = dW_t^{Q} - (\lambda_0 + \lambda_1 X_t)dt, \tag{2.6}$$

where $\lambda_0 = \left(\lambda_0^{US}, \lambda_0^{BR}\right)' \in \mathbb{R}^N$ and λ_1 is $N \times N$ matrix given by

$$\lambda_1 = \begin{bmatrix} \lambda_1^{US,US} & 0 \\ \lambda_1^{BR,US} & \lambda_1^{BR,BR} \end{bmatrix}.$$

As a result, the price P^{BR} of a defaultable bond is exponential affine, that is, $P^{BR}(t,T) = \exp\left(a^{BR}(\tau) + b^{BR}(\tau)X_t\right)$, where $\tau = T - t$, and a^{BR} and b^{BR} solve a system of Riccati differential equations:

$$\begin{aligned} b^{BR}(\tau)' &= -(\delta_1^r + \delta_1^s) - K^{\star\prime} b^{BR}(\tau) \\ a^{BR}(\tau)' &= -(\delta_0^r + \delta_0^s) + \xi^{\star\prime} K^{\star\prime} b^{BR}(\tau) + \frac{1}{2} b^{BR}(\tau)' \Sigma \Sigma' b^{BR}(\tau), \end{aligned} \tag{2.7}$$

with $K^\star = K + \Sigma\lambda_1$ and $\xi^\star = K^{\star -1}(K\xi - \Sigma\lambda_0)$. An explicit solution for this system of differential equations exists only in some special cases, such as diagonal K. However, the Runge-Kutta method provides accurate numerical approximations. Thus, the yield at time t with time to maturity τ is given by

$$\begin{aligned} Y_t^{BR}(\tau) = A^{BR}(\tau) + B_\theta^{BR,US}(\tau)\theta_t^{US} + \\ B_M^{BR,US}(\tau)M_t^{US} + B_M^{BR,BR}(\tau)M_t^{BR} + B_\theta^{BR,BR}(\tau)\theta_t^{BR}. \end{aligned} \tag{2.8}$$

If the loss given default rate is constant, i.e. $\ell_t = \ell$ for all t, then the term structure of default probabilities is given by (see Schönbucher, 2003):

$$\Pr(t,T) = 1 - E_t^{P}\left[\exp\left(-\int_t^{t+T} \frac{s_u}{\ell} du\right)\right], \tag{2.9}$$

which can be calculated similarly to the conditional expectation contained in the pricing equation, with the objective measure replacing the martingale measure. It turns out that $\Pr(t,\tau) = 1 - \exp(a^{Pr}(\tau) + b^{Pr}(\tau)X_t)$, where a^{Pr} and b^{Pr} are again solutions of Riccati differential equations:

$$\begin{aligned} b^{Pr\prime}(\tau) &= -\delta_1^s/\ell - K' b^{Pr}(\tau), \\ a^{Pr\prime}(\tau) &= -\delta_0^s/\ell + \xi' K' b^{Pr}(\tau) + \frac{1}{2} b^{Pr}(\tau)' \Sigma \Sigma' b^{Pr}(\tau). \end{aligned} \tag{2.10}$$

We close this section with two remarks. First, the reduced model can be replaced by a standard term structure model with macro factors: it suffices to let the US factors take the role of macro factors for the defaultable bonds. However, the interpretation of the spread as the instantaneous expected loss and the computation of model implied default probabilities are no longer possible. Second, all the models in this article are in the class of Gaussian models, the simplest specification of the affine family. The inclusion of macro variables and default substantially complicates the model and its estimation. Therefore, we decided not to use a model with stochastic volatility. However, note that macro factors such as the VIX volatility can approximately play the role of stochastic volatility of the non-Gaussian affine models. Furthermore, models with constant volatility are the best choice matching some stylized facts (as shown, for instance, by Duffee, 2002, and Dai and Singleton, 2002) and to describe corporate CDS spreads (see Berndt et al., 2004).

2.1. Impulse Response and Variance Decomposition. One way to evaluate the impact of macro shocks on the term structure of interest rates and default probabilities is through impulse response functions (IRF) and variance decompositions (VD).

The response of the yield $Y_t = A + BX_t$ is given by

(2.11)
$$\begin{array}{ccccc} B\Sigma\sqrt{\Delta t}\varepsilon_t & Be^{-K\Delta t}\Sigma\sqrt{\Delta t}\varepsilon_t & Be^{-2K\Delta t}\Sigma\sqrt{\Delta t}\varepsilon_t & Be^{-3K\Delta t}\Sigma\sqrt{\Delta t}\varepsilon_t & \ldots \\ t+0 & t+1 & t+2 & t+3 & \ldots \end{array}$$

and the response of the logarithm of the survival probability, $\log \Pr(t,\tau) = a^{\Pr} + b^{\Pr} X_t$, is

(2.12)
$$\begin{array}{ccccc} b^{\Pr}\Sigma\sqrt{\Delta t}\varepsilon_t & b^{\Pr}e^{-K\Delta t}\Sigma\sqrt{\Delta t}\varepsilon_t & b^{\Pr}e^{-2K\Delta t}\Sigma\sqrt{\Delta t}\varepsilon_t & b^{\Pr}e^{-3K\Delta t}\Sigma\sqrt{\Delta t}\varepsilon_t & \ldots \\ t+0 & t+1 & t+2 & t+3 & \ldots \end{array}$$

Moreover, the contribution corresponding to the j^{th} factor in the variance decomposition of $Y_{t+h}(\tau)$ and $\log \Pr_{t+h}(\tau)$ at time t are

$$\begin{aligned} VD_j(Y) &= B'(\tau)\left(\int_t^{t+h} e^{-K(t+h-u)}\Sigma_j\Sigma_j'(e^{-K(t+h-u)})'du\right)B(\tau), \\ VD_j(\log) &= b^{\Pr'}(\tau)\left(\int_t^{t+h} e^{-K(t+h-u)}\Sigma_j\Sigma_j'(e^{-K(t+h-u)})'du\right)b^{\Pr}(\tau). \end{aligned} \quad (2.13)$$

3. Data

Our sample consists of a daily series of the following variables: (i) constant maturity zero-coupon term structure of U.S. yields provided by the Federal Reserve (Fed); (ii) constant maturity zero-coupon term structure of Brazilian sovereign yields constructed by Bloomberg[9]; (iii) the implied volatility of S&P 500 index options measured by the Chicago Board Options Exchange Volatility Index - VIX; (iv) Brazilian Real/US Dollar exchange rate, (v) São Paulo Stock Exchange index - IBovespa[10], (vi) Brazilian domestic zero coupon yields extracted from DI x Pre swaps obtained from Brazilian Mercantile and Futures Exchange (BM&F)[11]. The first two data set are used as basic yields and the others take the role of observed (macro) factors in our model.

The sample begins on February 17, 1999, and ends on September 15, 2004, with a total of 1320 days. The sample starts one month after the change of the exchange rate regime from fixed to floating in January 1999, forced by a devaluation crisis. The maturities of the US and Brazilian sovereign yields are the same, namely 3 and 6 months, 1, 2, 3, 5, 7, 10, and 20 years while the maturities of the Brazilian domestic yields are 1, 3, and 36 months. Figure 1 depicts the US and Brazilian sovereign yields. Figure 2 shows the observed variables. Note that the American yield curve is almost flat in the beginning of the sample. After January 2001, short-yields decline over time and the shape of the term structure changes to upward sloping. In end of 2002, there is a stress movement in the Brazilian market due to a presidential succession process in which the candidate of the opposition won the election.

---

[9] The dataset of sovereign yields provided by Bloomberg is extracted from Brazilian Global bonds.

[10] Ibovespa is the main Brazilian stock market index.

[11] The ID rate is the average one-day interbank borrowing/lending rate, calculated by CETIP (Center of Custody and Financial Settlement of Securities) every business day. The ID rate is expressed in effective rate per annum, based on 252 business-days.

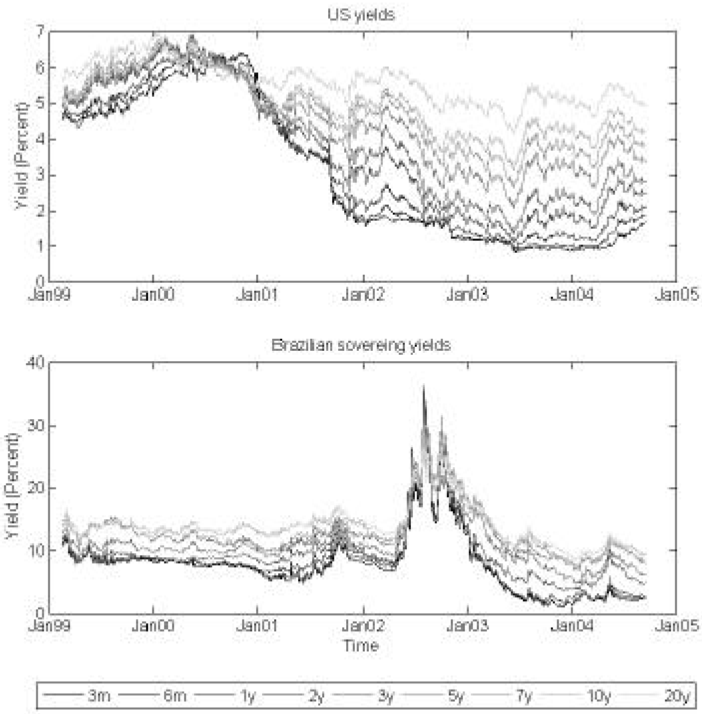


Figure 1. U.S. and Brazilian sovereign yields. This …gure contains time series of U.S. (top panel) and Brazilian sovereign (bottom panel) yields with time to maturity of 3 and 6 months, 1, 2, 3, 5, 7, 10 and 20 years between February 17, 1999 and September 15, 2004.

4. Estimation

The parameters are estimated via the maximum likelihood method. Although it is possible to make one-step estimations of the U.S. and Brazilian sovereign yield curves, it is computationally more interesting to work with a simpler technique using a two-step procedure, as in Du¢e et al (2003). We use the U.S. term structure as the reference curve (default-free curve) for our analysis. In the …rst step we estimate the reference curve using only latent factors. Then, conditional on the parameters and state vector of the U.S. curve we estimated the Brazilian sovereign yield curve.

We now describe the procedure adopted for a model with macro variables and default. The estimation of U.S. parameters is a particular case of this general framework. By stacking the parameters and state variables, the yield of a defaultable bond (Equation 2.8) can be written as

$$Y_t^{BR}(\tau) = A^{BR}(\tau) + B^{BR}(\tau)X_t, \tag{4.1}$$

where the dynamics of X_t is given by Equation 2.5.

The likelihood is the joint probability density function of the sequence of observed Brazilian sovereign yields $Y_t^{BR} = \left(Y_{t_1}^{BR}, \ldots, Y_{t_n}^{BR}\right)$ and macro factors M_t. It is possible to show that the transition density of $X_{t_i}|X_{t_{i-1}}$, denoted by f_X, is

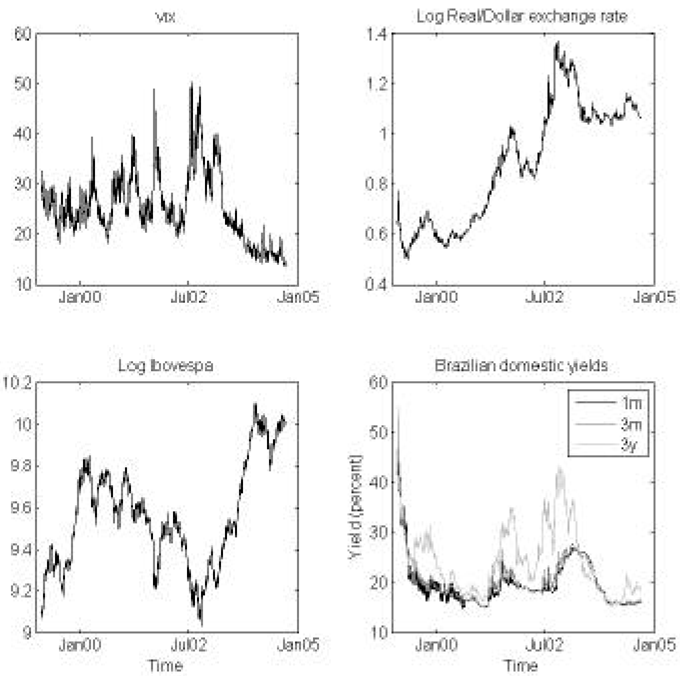


Figure 2. Observable variables. This figure contains time series of variables used as observable factors in our model between February 17, 1999 and September 15, 2004. The upper left panel shows the evolution of the VIX (implied volatility of S&P 500 index options). The upper right panel presents the logarithm of the Brazilian Real/US Dollar exchange rate. The left lower panel presents the logarithm of the Ibovespa (São Paulo Stock Exchange index), and the right lower panel shows the Brazilian domestic zero-coupon yields with time to maturity of 1, 3 and 36 months.

normally distributed with mean $\mu_i^{BR} = e^{-K(t_i - t_{i-1})} X_{t_{i-1}} + \left(I_N - e^{-K(t_i - t_{i-1})} \right) \xi$ and variance $(\sigma_i^{BR})^2 = \int_{t_{i-1}}^{t_i} e^{-K(t_i - u)} \Sigma \Sigma' e^{-K(t_i - u)'} du$ (see, for instance, Fackler, 2000).

Suppose first the vectors θ_t^{BR} and Y_t^{BR} have the same dimension, that is, we observe as many yields as latent variables. Then, we can invert a linear equation and find the unobserved factors θ_t^{BR} as a function of yields Y_t^{BR} and observable factors M_t^{BR}. Using change of variables, the log-likelihood function can be written as

$$L(Y_t, M_t, \psi) = \sum_{t=2} \log f_X (X_t | X_{t-1}, \psi) + (H - 1) \log \det |Jac|,$$

where H is the sample size, $\Theta = (\delta_0, \delta_1, K, \xi, \Sigma, \lambda_0, \lambda_1)$ is a vector stacking the model parameters, and the Jacobian matrix is

$$Jac = B^{BR}(\tau_1, \ldots, \tau_{N^{BR}}) = \begin{bmatrix} B^{BR}(\tau_1) \\ \vdots \\ B^{BR}(\tau_{N^{BR}}) \end{bmatrix}, \tag{4.2}$$

where $\tau_1, \ldots, \tau_{N^{BR}}$ are the time to maturities of the observable Brazilian yields.

If we want to use additional yields, direct inversion is not possible. This is known as "stochastic singularity". One solution is to follow Chen and Scott (1993), and add measurement errors to the extra yields. Let N^{BR}_{obs} be the number of Brazilian sovereign yields observed on each day, $N^{BR}_{obs} > N^{BR}$ where N^{BR} is the size of X^{BR}_t. We select N^{BR} yields to be priced without error. The other $(N^{BR}_{obs} - N^{BR})$ are priced with independent normal measurement errors. Therefore, the log-likelihood function is

$$L(Y_t, M_t, \Theta) = \sum_{t=2} \log f_X(X_t | X_{t-1}, \Theta) + (H - 1) \log |\det Jac| + \frac{1}{2} \sum_{t=2} u_t' \Omega^{-1} u_t,$$

where u_t is the vector of yield measurement errors and Ω represents the covariance matrix for u_t, estimated using the sample covariance matrix of the u_t's implied by the extracted state vector, and $Jac = B^{BR}(\tau_1, \cdots, \tau_{N^{BR}_{obs}})$.

In order to complete the estimation procedure, it is necessary to identify the model. If the model is sub-identified then there are more than one set of parameters that generate the same likelihood. Therefore, not all parameters can be estimated. On the other hand over-identified models produce sub-optimal results that may distort the impulse response functions. However, identification of parameters in a state-space system is tricky. In the next subsection, we provide identification strategies for some specifications of our model based on the results of Dai and Singleton (2000).

4.1. **Model Identification.** Here, we show how to identify the parameters of a Gaussian affine model with macro factors and credit spreads. This approach is based on the study of Dai and Singleton (2000).

First we consider the default-free case. Suppose there are p macro variables M and q latent variables θ. The vector $X = (M, \theta)$ follows a Gaussian affine dynamics:

$$\begin{aligned} dX_t &= \begin{bmatrix} dM_t \\ d\theta_t \end{bmatrix} = \left(\begin{bmatrix} K_{M,M} & K_{M,\theta} \\ K_{\theta,M} & K_{\theta,\theta} \end{bmatrix} \left(\begin{bmatrix} \xi_M \\ \xi_\theta \end{bmatrix} - \begin{bmatrix} M_t \\ \theta_t \end{bmatrix} \right) \right) dt \\ &+ \begin{bmatrix} \Sigma_{M,M} & \Sigma_{M,\theta} \\ \Sigma_{\theta,M} & \Sigma_{\theta,\theta} \end{bmatrix} \begin{bmatrix} dW^P_M(t) \\ dW^P_\theta(t) \end{bmatrix} = K(\xi - X_t)dt + \Sigma dW^P(t). \end{aligned} \tag{4.3}$$

The instantaneous short-term rate is given by $r_t = \delta_0 + \delta_1 \cdot X_t$ while the market price of risk obeys Equation 2.6. Hence, the dynamics of X in the risk-neutral measure is $dX = K^*(\xi^* - X_t)dt + \Sigma dW^Q(t)$ and the yield curve is an affine function of X, $Y_t(\tau) = A(\tau) + B^M(\tau)M_t + B^\theta_t(\tau)\theta = A(\tau) + B(\tau)X_t$. The parameter vector is denoted by $\Theta = (\delta_0, \delta_1, K, \xi, \lambda_0, \lambda_1, \Sigma)$.

Some of the above parameters must be arbitrarily fixed, otherwise there are multiple solutions to the estimation problem since we can define operators that preserve the likelihood as shown below.

Let $L \in \mathbb{R}^{(p+q)\times(p+q)}$ be a non-singular matrix and $v \in \mathbb{R}^{p+q}$ a vector such that

$$L = \begin{pmatrix} I & 0 \\ \alpha & \beta \end{pmatrix} \text{ and } v = \begin{pmatrix} 0 \\ v_\theta \end{pmatrix},$$

where $I \in \mathbb{R}^{p\times p}$ is the identity matrix, $\alpha \in \mathbb{R}^{q\times p}$, $\beta \in \mathbb{R}^{q\times q}$, and $v^\theta \in \mathbb{R}^q$. Consider the following maps:

(4.4)

$$T_{L,v}\{\Theta, X\} = \{(\delta_0 - \delta_1' L^{-1}v, (L')^{-1}\delta_1, LKL^{-1}, v + L\xi, \lambda_0 - \lambda_1 L^{-1}v, \lambda_1 L^{-1}, L\Sigma), LX + v\}$$

and

$$T_O\{\Theta, X\} = \{(\delta_0, \delta_1, K, \xi, \lambda_0, \lambda_1, \Sigma O'), X\}, \tag{4.5}$$

where $O \in \mathbb{R}^{(p+q)\times(p+q)}$ is a rotation matrix.

Proposition 1. The operators $T_{L,v}$ and T_O preserve the likelihood of the affine model defined above under the Chen-Scott (1993) estimation procedure.

Proof

The log-likelihood $\mathcal{L}$ of the affine model under the Chen-Scott (1993) inversion is

$$\begin{gathered}\mathcal{L}(\Theta, X) = \log f_Y(Y_{t_1}, ..., Y_{t_H} | \Theta, X) = \\ \log f_X(X_{t_1}, ..., X_{t_H} | \Theta) + \log f_u(u_{t_1,...,}u_{t_H}) + \log|\det Jac|^{H-1} = \\ (H-1)\log|\det \hat{B}^\theta| + \sum_{t=2}^{H} \log f_{X_t|X_{t-1}}(X_t|\Theta) + \log f_u(u_t) = \\ (H-1)\log|\det \hat{B}^\theta| - \tfrac{1}{2}(H-1)\log\det\left[\Delta t \left(e^{-K\Delta t}\right)\Sigma\Sigma'\left(e^{-K\Delta t}\right)'\right] \\ + \sum_{t=2}^{H}\log f_u(u_t) - \tfrac{1}{2}\sum_{t=2}^{H}(X_t - \mu)'\left[\Delta t\left(e^{-K\Delta t}\right)\Sigma\Sigma'\left(e^{-K\Delta t}\right)'\right]^{-1}(X_t - \mu),\end{gathered}$$

where $\mu = e^{-K\Delta t}\xi + (1 - e^{-K\Delta t}\xi)X_{t-1}$, $\Delta t = t_i - t_{i-1}\ \forall i$, H is the sample size, and $\hat{B}^\theta(\cdot)$ is evaluated at the time to maturities of yields without measurement errors (see Equation 4.2).

We begin by proving that $\mathcal{L}(\Theta, X) = \mathcal{L}(T_{L,v}(\Theta, X))$. The strategy of the proof is to analyze what happens with each of the four terms of the log-likelihood when the operator $T_{L,v}$ is applied. First, note that the expression under the last summation symbol is preserved. The transformation of μ is

$$\begin{gathered}\mu(T_{L,v}(\Theta, X)) = e^{-LKL^{-1}\Delta t}L\xi + \left(1 - e^{-LKL^{-1}\Delta t}\right)LX_{t-1} = \\ Le^{-K\Delta t}L^{-1}L\xi + (1 - Le^{-K\Delta t}L^{-1})LX_{t-1} = L\mu.\end{gathered}$$

Then, applying $T_{L,v}$ on the last summation expression of the log-likelihood, we have

$$\begin{gathered}(LX_t - L\mu)'\left[\left(e^{-LKL^{-1}\Delta t}L\Sigma\sqrt{\Delta t}\right)\left(e^{-LKL^{-1}\Delta t}L\Sigma\sqrt{\Delta t}\right)'\right]^{-1}(LX_t - L\mu) = \\ (X_t - \mu)'L'\left[\left(Le^{-K\Delta t}L^{-1}L\Sigma\sqrt{\Delta t}\right)\left(Le^{-K\Delta t}L^{-1}L\Sigma\sqrt{\Delta t}\right)'\right]^{-1}L(X_t - \mu) = \\ (X_t - \mu)'L'\left[L'^{-1}\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)'^{-1}\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)^{-1}L^{-1}\right]L(X_t - \mu) = \\ (X_t - \mu)'\left[\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)'^{-1}\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)^{-1}\right](X_t - \mu).\end{gathered}$$

The second term of the log-likelihood changes to

$$-\frac{1}{2}(H-1)\log\det\left[\left(e^{-LKL^{-1}\Delta t}L\Sigma\sqrt{\Delta t}\right)\left(e^{-LKL^{-1}\Delta t}L\Sigma\sqrt{\Delta t}\right)'\right] =$$

$$-\frac{1}{2}(H-1)\left\{\log\det\left[\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)'\right]+2\log\det L\right\} = \tag{4.6}$$

$$-\frac{1}{2}(H-1)\left[\log\det\left[\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)'\right]\right]-(H-1)\log\det L.$$

It is easy to see that

$$\begin{aligned}(H-1)\log|\det\Psi^{\theta}|(T_{L,v}(\Theta,X)) &= (H-1)\log|\det\beta^{-1}\Psi^{\theta}|\\ &= (H-1)\log|\det\Psi^{\theta}|+(H-1)\log|\det\beta^{-1}|.\end{aligned}$$

Since $\det L=\det\beta$, the last term that appeared in (4.6) cancels out with the last term in the expression above.

Moreover, it is also easy to see that u_t does not change under the transformation $T_{L,v}$.

Finally, $\mathcal{L}(\Theta,X)=\mathcal{L}(T_O(\Theta,X))$ since the only expression affected by the rotation is preserved:

$$\begin{aligned}&\left[\left(e^{-K\Delta t}\Sigma O'\sqrt{\Delta t}\right)\left(e^{-K\Delta t}\Sigma O'\sqrt{\Delta t}\right)'\right]^{-1}\\ &=\left[\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)\left(e^{-K\Delta t}\Sigma\sqrt{\Delta t}\right)'\right]^{-1}.\end{aligned}$$

□

Therefore, there are infinite parameter vectors with the same likelihood. Hence, before estimation through the maximum likelihood method, some parameters must be fixed. On the other hand, the imposition of over-identifying restrictions may produce sub-optimal results that distort the impulse response functions. The model can be considered identified if all the degrees of freedom of the model, which are given by α, β, v_θ and O, are eliminated.

Note that v_θ can always be used to set $\xi_\theta=0$. In addition, the rotation O implies that Σ must be a triangular matrix for a given state vector order. Hence, we choose $\Sigma_{\theta,\theta}$ and $\Sigma_{M,M}$ to be lower triangular and $\Sigma_{M,\theta}=0$. Finally, α and β can be set so that $\Sigma_{\theta,\theta}=I$, $\Sigma_{M,\theta}=0$, and $K_{\theta,\theta}$ is lower triangular. This completes the identification of the default-free case.

We now turn to the case with default. Formally speaking, the reduced credit risk model of Duffie and Singleton (1999) is simply a higher-dimensional affine model and the same identification procedure can be applied. There are, however, two subtleties involved.

The first is that there are natural restrictions that can be placed to the default model coming from economic considerations. For instance, we have considered that the American yield curve and macro factors affect the Brazilian curve, but not vice versa. However, the model must be first identified from the econometric point of view before additional restrictions are imposed, otherwise the same parameters might be fixed twice, leaving unresolved degrees of freedom.

The second point is that in the default-free case was illustrated supposing that the macro factors are "more endogenous" than the latent factors. In the default case, $X=(\theta^{US},M^{US},M^{BR},\theta^{BR})$, thus the American latent factors come before the Brazilian factors, which would in principle change the operator $T_{L,v}$ and consequently the degrees of freedom. The other inversion, namely the American macro vector coming after the latent vector, is due to the fact that only the VIX is considered and it does not interfere with the identification procedure.

However, since we use a two-step procedure, the parameters and state factors related to the American term structure are estimated first. So, we can think of the American latent factors as if they were "macro" factors and proceed to the identification considering that $\tilde{M}^{BR} = (\theta^{US}, M^{US}, M^{BR})$ is in fact the macro vector for the default case.

In summary, the economic restrictions impose that $\delta_1^r = (\delta_1^{r,US}, 0)$ and that the matrix K is block-triangular, which means that Brazilian factors do not affect American factors. Therefore the identified Σ is given by:

$$\begin{pmatrix} \tilde{\Sigma}_{MM} & 0 \\ 0 & I \end{pmatrix}, \text{ where } \tilde{\Sigma}_{MM} = \begin{pmatrix} I & 0 & 0 \\ 0 & \Sigma^{US,US}_{M,M} & 0 \\ \Sigma^{BR,US}_{M,\theta} & \Sigma^{BR,US}_{M,M} & \Sigma^{BR,BR}_{M,M} \end{pmatrix}.$$

5. Results

In this section we analyze the results of three different specifications of our model estimated by the maximum likelihood method described in Section 4. We begin with a simple macro-to-yield without default specification. In order to avoid local maxima, many trial numerical optimizations are performed using the Nelder-Mead Simplex algorithm until stable results are obtained. Then, taking advantage of these results, we select starting vectors for the estimation of two higher dimensional models with default. After that, other independent trial maximization starting from random vectors are performed. Finally we choose the best results. Although this procedure may be path-dependent, the "curse of dimensionality" does not allow the use of a complete grid of random starting points as would be desirable.

5.1. Macro-to-yield without default.

The simplest specification of our model is characterized by a macro-to-yield dynamics without default. It is exactly the model of Ang and Piazzesi (2003) applied to the Brazilian yield curve. The absence of default implies that American latent factors (θ^{US}) are unnecessary. In a macro-to-yield model the observable factors affect the latent factors but not vice versa. This means that $K^{BR,BR}_{M,\theta}$ is a matrix of zeros.

The macro-to-yield without default specification presents three state variables, $X = (M, \theta_1^{BR}, \theta_2^{BR})$. It serves to indicate the relevant macro factors for the sovereign yield curve, which are then selected for use in the other models. To extract Brazilian latent factors, we set the 3-month and 5-year sovereign yields to be flawless. Nine versions are estimated, each having a different observed factor M: (1) VIX; (2) logarithm of the Brazilian Real/US Dollar exchange rate (LEX); (3) logarithm of the IBovespa (LIBOV); (4) BM&F 1-month yield (B1m); (5) BM&F 3-year yield (B3y); (6) BM&F slope (Bsl) = B3y - B1m; (7) Fed 1-month yield (F1m); (8) Fed 10-year yield (F10y), and (9) Fed slope (Fsl)= F10y - F1m.

Table 1 presents the log-likelihood divided by the number of observations (L/H) and the mean (for the nine maturities) of the absolute measurement errors in basis points (MAE) for all specifications. These measures can be used to evaluate the different versions of a model. Table 1 also presents the correlations between factor

1 (θ_1^{BR}) and the slope of the Brazilian sovereign term structure ($\rho_{1,s}$) and between factor 2 (θ_2^{BR}) and the level of the Brazilian sovereign term structure ($\rho_{2,l}$). The likelihood does not vary significantly, but the specifications that included US rates show slightly higher values. The mean absolute measurement error is around 60 basis points. The latent factor θ_2^{BR} represents the level, since it is highly correlated with this factor in all cases, while θ_1^{BR} can be interpreted as the slope due to its positive correlation with the slope of the yield curve.

Table 1. Summary of results of the macro-to-yield without default model.

	VIX	LEX	LIBOV	B1m	B3y	Bsl	F1m	F10y	Fsl
L/H	44.7	44.3	44.3	44.8	44.8	45.0	47.5	47.5	47.1
$\rho_{1,s}$	0.20	0.37	0.29	0.59	0.56	0.57	0.66	0.69	0.61
$\rho_{2,l}$	0.99	0.83	0.98	0.94	0.94	1.00	0.94	0.84	0.94
MAE	54	66	56	62	62	58	62	61	62

This table presents the log-likelihood divided by the number of observations (L/H), the mean (for the nine maturities) of the absolute measurement errors in basis points (MAE), and the correlations between factor 1 (θ_1^{BR}) and the slope of the Brazilian sovereign yield curve ($\rho_{1,s}$) and between factor 2 (θ_2^{BR}) and the level of the Brazilian sovereign yield curve ($\rho_{2,l}$). The macro-to-yield without default model presents only one observable factor in each specification. They are (1) VIX; (2) logarithm of the BR Real/US Dollar exchange rate (LEX); (3) logarithm of the Ibovespa (LIBOV); (4) BM&F 1-month yield (B1m); (5) BM&F 3-years yield (B3y); (6) BM&F slope (Bsl) = B3y - B1m; (7) Fed 1-month yield (F1m); (8) Fed 10-years yield (F10y), and (9) Fed slope (Fsl)= F10y - F1m.

In order to measure the relative contributions of the macro and latent factors to forecast variances we perform variance decompositions. Table 2 presents the proportion of the 1-month and 9-month ahead forecast variance of the {3m, 3y, 20y}-yields attributable to each observable factor used in each of the nine versions. This provides a comparison of the importance of the different macro variables for the sovereign yield curve by showing the macro participation on the variance of the yields 1 and 9 months after the shock. The order of the impact can be summarized as follows: VIX and BM&F slope present the largest effect, accounting for up to 69% and 79% of the 20-year yields 9 months after the shock. Although still significant, the contribution of Brazilian Real/US Dollar exchange rate, 10-years Fed yield, Fed slope, and IBovespa are much smaller. Finally, BM&F 1-month and 3-year yields, and Fed 1-month yield show negligible effect.

5.2. Macro-to-yield with default. In this subsection, we introduce default risk into the previous specification. Again, we assume that the state variables follow a macro-to-yield dynamics. There is a need for another latent factor besides the macro factor and the two Brazilian latent factors. The job of this new factor is to capture the US term structure, which represents the reference curve. The parameters corresponding to the US latent factor are estimated in a first step, while the other parameters are estimated conditional on the first step. The American latent factor is obtained from the yield with 3 months maturity while the Brazilian

Table 2. Variance decompositions of the macro-to-yield without default model.

Yields	1 month ahead								
	VIX	LEX	LIBOV	B1m	B3y	Bsl	F1m	F10y	Fsl
3m	15	7	1	0	0	16	0	0	0
3y	23	9	0	0	0	23	0	4	0
20y	54	9	6	0	0	50	0	8	0
Yields	9 months ahead								
	VIX	LEX	LIBOV	B1m	B3y	Bsl	F1m	F10y	Fsl
3m	31	7	22	0	0	46	0	0	6
3y	46	11	13	0	0	61	0	10	7
20y	69	14	21	0	0	79	0	16	7

This table presents the proportion (in percent) of the one month and nine months ahead forecast variance of the {3m, 3y, 20y}-yields attributable to each observable factor. The macro-to-yield without default model presents only one observable factor in each speci...-cation. They are (1) VIX; (2) logarithm of the BR Real/US Dollar exchange rate (LEX); (3) logarithm of the Ibovespa (LIBOV); (4) BM&F 1-month yield (B1m); (5) BM&F 3-years yield (B3y); (6) BM&F slope (Bsl) = B3y - B1m; (7) Fed 1-month yield (F1m); (8) Fed 10-years yield (F10y), and (9) Fed slope (Fsl)= F10y - F1m.

latent factors are obtained from the sovereign yields with maturities of three months and ...ve years.

Table 3. Summary of results of the macro-to-yield with default model.

	y.o.	VIX	Bsl	Fsl	B3y
L/H	42.1	47.0	48.2	49.9	47.8
$\rho_{1,s}$	0.28	0.37	0.42	-0.17	0.19
$\rho_{2,l}$	0.99	0.97	0.98	0.96	0.86
MAE	68	53	50	59	54

This table presents the log-likelihood divided by the number of observations (L/H), the mean (for the nine maturities) of the absolute measurement errors in basis points (MAE), and the correlations between factor 1 (θ_1^{BR}) and the slope of the Brazilian sovereign yield curve ($\rho_{1,s}$) and between factor 2 (θ_2^{BR}) and the level of the Brazilian sovereign yield curve ($\rho_{2,l}$). The macro-to-yield with default model presents one observable factor, one latent factor driving the US curve and two latent factors driving the Brazilian curve. The observable factors are (1) VIX; (2) BM&F slope (Bsl) = B3y - B1m, (3) Fed slope (Fsl)= F10y - F1m, and (4) BM&F 3-years yield (B3y). The y.o. model refers to a speci...cation in which only yields are used, that is, a speci...cation without observable factors.

In view of the results of the previous subsection, we divide the observable factors into three groups. The ...rst one is composed of the VIX and BM&F slope which are the factors that have the largest impact on the yields. The intermediate group consists of the Brazilian Real/US Dollar exchange rate, 10-year Fed yield, Fed slope, and IBovespa. The third group presents little effect on yields, being formed of BM&F 1-month and 3-year yields, and Fed 1-month yield. In order to understand the impact of macro variables on the yields in a model with default, we use both factors of the first group, one factor of the second group (Fed slope), and one factor of the third group (BM&F 3-years yield)[12].

Table 4. Variance decompositions of the yields of the macro-to-yield with default model.

Model		y.o.		VIX		Bsl		Fsl		B3y	
Factor	Yield	1m	9m	1m	9m	1m	9m	1m	9m	1m	9m
θ^{US}	3m	0	1	0	4	0	1	0	6	1	4
	3y	0	0	0	2	0	0	0	5	1	2
	20y	0	0	0	1	0	0	0	5	1	2
Macro	3m	-	-	15	37	1	7	2	14	1	11
	3y	-	-	25	50	2	8	1	16	0	2
	20y	-	-	56	70	5	9	1	17	4	3
θ_1^{BR}	3m	13	26	10	8	25	14	51	23	30	32
	3y	1	20	1	2	8	11	24	17	79	89
	20y	18	12	9	5	7	11	2	13	85	91
θ_2^{BR}	3m	87	74	75	51	74	77	47	57	68	53
	3y	99	80	74	47	89	80	75	63	20	7
	20y	82	88	35	24	88	80	96	65	10	4

This table presents the proportion (in percent) of the one month and nine months ahead forecast variance of the {3m, 3y, 20y}-yields attributable to each observable factor in the macro-to-yield with default model. The macro-to-yield with default model presents one observable factor, one latent factor driving the US curve and two latent factors driving the Brazilian curve. The observable factors are (1) VIX; (2) BM&F slope (Bsl) = B3y - B1m, (3) Fed slope (Fsl)= F10y - F1m, and (4) BM&F 3-years yield (B3y). The y.o. model refers to a specification in which only yields are used, that is, a specification without observable factors.

Table 3 summarizes the results of some versions of the macro-to-yield with default model. It shows the likelihood, correlations and measurement errors of the yields of each specification. The first column refers to the yields only model (y.o.) in which only latent factors are used. The others are macro-to-yield models with VIX, BM&F slope, Fed slope, and BM&F 3-years yield as observable factors. The inclusion of the U.S. reference curve produces a gain in likelihood and in fit, because the measurement errors are lower. The latent factor θ_2 remains highly linked to the level of the sovereign yields.

Table 4 presents the variance decomposition of the {3m, 3y, 20y}-yield for one and nine months ahead. We see that the VIX is still very important, contributing with up to 70% of the 20-years yield variation. Other variables accounted for less, but still some effect can be attributed to them. Furthermore, in the y.o. version the US factor seems to be insignificant.

[12]Models with other observable factors from the second and third groups were also tested providing similar qualitative results.

Table 5. Variance decompositions of the default probabilities of the macro-to-yield with default model.

Model		y.o.		VIX		Bsl		Fsl		B3y	
Factor	Term	1m	9m	1m	9m	1m	9m	1m	9m	1m	9m
θ^{US}	3m	0	0	0	1	0	0	1	4	0	1
	3y	0	1	19	60	0	0	4	4	12	25
	20y	0	1	69	92	5	6	3	3	51	65
Macro	3m	-	-	54	61	5	9	4	18	8	25
	3y	-	-	48	26	9	10	20	22	10	26
	20y	-	-	19	5	9	9	22	23	9	9
θ_1^{BR}	3m	22	28	9	7	15	12	34	19	12	21
	3y	27	29	6	3	12	11	17	15	11	18
	20y	27	29	2	1	11	11	15	15	11	24
θ_2^{BR}	3m	78	72	37	30	80	79	62	60	70	56
	3y	73	71	27	11	79	78	60	59	45	31
	20y	73	71	10	02	74	74	59	59	16	6

This table lists the contribution (in percent) of each factor to the one month and nine months ahead forecast of the {3m, 3y, 20y} default probabilities within the macro-to-yield with default model.

We also calculate the variance decompositions of the logarithm of the default probabilities, which can be seen in Table 5. All results presented in this paper are obtained using a …xed loss given default $\ell = 50\%$. This particular choice is, of course, arbitrary, however there is empirical evidence that the mean of the loss rate is around this value (see, for instance, Moody's, 2008)[13]. The VIX is responsible for the greatest e¤ect, especially in the short-term. According to the model, in the 1- and 9-month horizon, VIX accounts, respectively, for 54% and 61% of the 3-month default probability. The BM&F and Fed slopes and BM&F 3-year yield explain 5%, 4% and 8% for 1-month ahead, and 9%, 18% and 25% for 9-month ahead, respectively, of the 3-month default probability. On the other hand, the Fed slope has the highest explanatory power for long-term default probability among the macro factors.

5.3. Bilateral models. In this subsection we present our main model. It has one American latent factor, one American macro factor (VIX), one Brazilian macro factor and two Brazilian latent factors. The Brazilian macro factor has a bilateral interaction with the Brazilian sovereign factors, that is, the macro factors and the sovereign yield curves fully interact. This means that $K_{M,\theta}^{BR,BR} \neq 0$. Once more, the American latent factor is obtained from the yield with maturity of three months while Brazilian latent factors are extracted considering that sovereign yields with maturities of three months and …ve years are priced without error.

We …x VIX as the American macro factor since it presents the best explanatory power for the simpler models analyzed in the previous subsections. We test four

[13] In order to verify the sensitivity of the results to the loss rate, we tested other values ($\ell = 25\%$ and $\ell = 75\%$) in the macro-to-yield with default model. From a qualitative point of view the results were very similar.

Table 6. Summary of results of bilateral model with default.

	VIX Bsl	VIX LIBOV-EX	VIX B3m	VIX B3y
L$/H$	52.5	55.6	52.9	52.9
$\rho_{1,s}$	0.48	0.86	0.08	0.04
$\rho_{2,l}$	0.93	0.96	0.92	0.93
MAE	47	51	46	46

This table presents the log-likelihood divided by the number of observations (L$/H$), the mean (for the nine maturities) of the absolute measurement errors in basis points (MAE), and the correlations between factor 1 (θ_1^{BR}) and the slope of the Brazilian sovereign yield curve ($\rho_{1,s}$) and between factor 2 (θ_2^{BR}) and the level of the Brazilian sovereign yield curve ($\rho_{2,l}$). The bilateral model with default presents one observable American factor (VIX), one latent factor driving the US curve, two latent factors driving the Brazilian curve and one observable Brazilian factor. The observable Brazilian factors are (1) BM&F slope (Bsl); (2) logarithm of the Bovespa index in US dollar (LIBOV-EX) (3) BM&F 3-months yield (B3m), and (4) BM&F 3-years yield (B3y).

speci…cations, which only di¤er with respect to the Brazilian macro factor. The …rst speci…cation takes the BM&F slope as the Brazilian macro factor. This is a very natural choice because this slope is the observable Brazilian factor that best explains the yield variations according to the macro-to-yields models. The second use the logarithm of the IBovespa in U.S. Dollars. This variable combines in single factor the information of two sources of uncertainty that present fairly good explanatory power in the macro-to-yield without default framework. Finally, although Brazilian domestic yields present little e¤ect, we consider the 3-month and 3-year Brazilian yields as domestic factors just to implement a robustness test.

Table 6 contains statistical measures of some versions of the bilateral model. Their likelihoods have increased in relation to the previous models, which indicates that the second macro factor and the bilateral dynamics add information and improve the in-sample …t, with the speci…cation containing the IBovespa presenting slightly higher likelihood. Also, the mean measurement errors of yields decreased to about 50 basis points. The unobservable factor θ_2 can still be interpreted as the level of the sovereign curve, but θ_1 is in some cases uncorrelated to the slope.

Table 7 reports the variance decomposition of {1m, 3y, 20y}-yields for forecast horizons of one and nine months ahead. In line with the preliminary models, the VIX is again the most important macro factor in‡uencing the yields. The e¤ect is stronger on the long end of the curve. Among the domestic variables, only the BM&F slope presents signi…cant explanatory power. Note that the latent factor related with the level of the sovereign curve is responsible for a large amount of yield variations. This suggests the existence of idiosyncratic sources of uncertainty in the sovereign yield curve that are not explained by the observable factors used in our model. This result is in agreement with the …ndings of Ang and Piazzesi (2003) and Diebold et al. (2006).

Table 8 presents the variance decomposition of the default probabilities. We now analyze in more details the 9-month horizon decomposition, since in this case the e¤ect of the initial condition is attenuated. Note that in all speci…cations, the US latent factor (approximately the Fed short rate) shows almost no e¤ect on

Table 7. Variance decompositions of the yields of the bilateral model with default.

Model		VIX Bsl		VIX LIBOV-EX		VIX B3m		VIX B3y	
Factor	Yield	1m	9m	1m	9m	1m	9m	1m	9m
θ^{US}	3m	0	2	1	3	1	2	1	3
	3y	0	1	0	5	0	1	0	1
	20y	0	0	0	4	0	0	0	0
M^{US}	3m	2	20	2	21	4	26	4	32
	3y	2	23	4	21	5	33	5	39
	20y	27	38	36	16	20	48	20	53
M^{BR}	3m	2	7	0	0	0	2	1	1
	3y	3	10	0	1	0	2	0	1
	20y	3	9	0	0	0	1	1	2
θ_1^{BR}	3m	38	22	68	50	76	12	17	9
	3y	13	9	19	52	3	4	2	1
	20y	2	4	5	70	0	1	0	1
θ_2^{BR}	3m	58	48	29	25	19	57	77	56
	3y	81	57	76	21	91	61	93	58
	20y	67	49	59	9	80	50	79	43

This table presents the proportion (in percent) of the one month and 9 months ahead forecast variance of the {3m, 3y, 20y}-yields attributable to each observable factor in the bilateral model with default. The bilateral model with default presents one observable American factor (M^{US} = VIX), one latent factor driving the US curve, two latent factors driving the Brazilian curve and one observable Brazilian factor (M^{BR}). The observable Brazilian factors are (1) BM&F slope (Bsl); (2) logarithm of the Bovespa index in US dollar (LIBOV-EX) (3) BM&F 3-months yield (B3m), and (4) BM&F 3-years yield (B3y).

short-term default probabilities. However, for the long-term (20 years), it is the principal factor, explaining around 80% of changes of implied default probabilities nine months ahead. The e¤ect of the VIX is smaller over the long-term, but about 50% of changes in implied short-term default probabilities are attributable to changes in this observable factor. Among the domestic factors, only the slope of the Brazilian local term structure has a relatively important e¤ect, accounting for 11% of changes in implied short-term default probabilities. Thus, we can conclude that, given our model and sample, the domestic rates, and also the IBovespa are not relevant sources driving default probability movements.

Figure 3 compares the evolution of the 1-year survival probabilities (one minus default probabilities) over the sample period. It can be seen that changing the domestic macro factor does not signi…cantly alter the probabilities. Observe that all versions capture the Brazilian electoral crisis in the second half of 2002, with the y.o. model having the largest impact on survival probability. The 1-year ahead survival probabilities fell from an average of 85% to around 70%, recovering later to around 90%.

In order to gauge the response of yields due to an unexpected change in state variables, we calculate impulse response functions. Figures 4, 5 and 6 show the

Table 8. Variance decompositions of the default probabilities of bilateral model with default.

Model		VIX Bsl		VIX LIBOV-EX		VIX B3m		VIX B3y	
Factor	Term	1m	9m	1m	9m	1m	9m	1m	9m
θ^{US}	3m	0	0	2	2	0	0	0	0
	3y	7	21	21	59	8	29	9	29
	20y	47	73	80	93	51	79	52	78
M^{US}	3m	19	32	16	31	29	42	34	56
	3y	31	31	26	18	36	33	55	51
	20y	18	11	7	3	20	10	29	16
M^{BR}	3m	8	11	0	0	1	3	1	1
	3y	11	10	1	1	3	3	2	2
	20y	6	3	0	0	1	1	1	1
θ_1^{BR}	3m	21	12	61	49	11	9	6	3
	3y	11	6	38	16	9	7	1	0
	20y	6	2	10	3	5	2	0	0
θ_2^{BR}	3m	52	44	22	18	59	46	59	39
	3y	41	31	14	6	43	29	33	17
	20y	24	11	3	1	23	8	18	5

This table lists the contribution (in percent) of each factor to the the one month and nine months ahead forecast of the {3m, 3y, 20y} default probabilities within the bilateral model with default.

e¤ect of a shock to US latent factor (θ_1^{US}), VIX and observable domestic factors (BM&F yields and slope and IBovespa in US Dollars), respectively, on the Brazilian {3m, 3y, 20y}-yields up to 18-months after the shock. The size of the shock is one standard deviation of a monthly variation of a state variable. In the next three months after a shock on VIX, yields rise about 1% and then fall. Changes in either the domestic short or long rate do not result in changes of the sovereign yields. The same is true for the domestic stock exchange index (IBovespa). However, a positive BM&F slope shock causes an increase in the yields. This may indicate a change of expectations of a future rise in in‡ation.

We now turn to survival probabilities. Figures 7, 8, 9 show the impact of a one deviation increase of a monthly variation of the US latent factor, VIX and observable domestic factors, respectively, on the survival probabilities in the next three months, three years and twenty years. It shows that the survival probability falls by up to 4% in relative terms due to a shock in the Fed short rate. An increase in VIX also decreases the survival probability about 1.5% in relative terms. Among the domestic factors, only the BM&F slope has some impact, decreasing the long-term survival probability by about 0.7% in relative terms.

6. Conclusion

We proposed a model that combines an a¢ ne yield dynamics with macro factors and credit risk. The model was estimated in two steps using the US and Brazilian sovereign yield curves. The credit spreads, the macro factors and the US yield curve

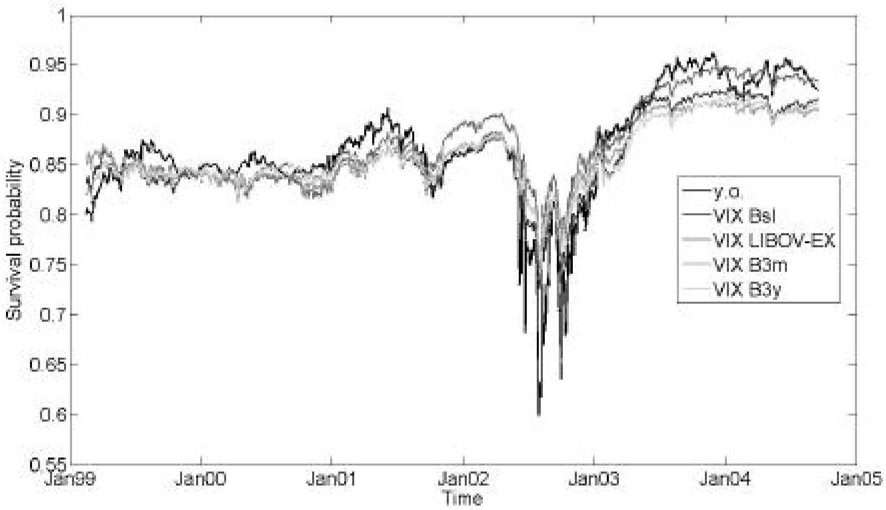


Figure 3. Survival probabilities. This figure shows the 1-year survival probabilities extracted from some versions of the bilateral model and from y.o. model between February 17, 1999 and September 15, 2004. The bilateral model presents one observable American factor (VIX), one latent factor driving the US curve, two latent factors driving the Brazilian curve and one observable Brazilian factor. The observable Brazilian factors are (1) BM&F slope (Bsl); (2) logarithm of the IBovespa in U.S. dollars (LIBOV-EX) (3) BM&F 3-month yield (B3m), and (4) BM&F 3-year yield (B3y). The y.o. model refers to a specification in which only yields are used, that is, a specification without observable factors.

have contemporaneous and lagged interaction. We were able to test how selected domestic and external macro factors such as the Brazilian Real/US Dollar exchange rate, VIX (volatility index of S&P 500), IBovespa (São Paulo stock exchange index) and domestic yield curve influence the spreads and default probabilities. The model was identified before making restrictions motivated by economic assumptions. Our findings indicate that the VIX and U.S. yield curve are the most important factors driving the Brazilian sovereign term structure and default probabilities. This result is consistent with the fact that credit risk premia of sovereign bond are highly correlated with the US economic conditions. The VIX has a high impact on 20-year bond yields and on short-term default probabilities, while the fed fund rate has high explanatory power on the long-term default probabilities. Among the domestic factors, only the slope of the local yield curve shows a significant effect on the Brazilian credit spread. However, a significant portion of variations in yields and default probabilities are explained by an unobservable factor highly correlated with the level of the Brazilian sovereign curve. Due to lack of an extensive historical dataset, we estimated a continuous-time version with daily observations, which limited the choices of macro variables. Future work can test monthly models,

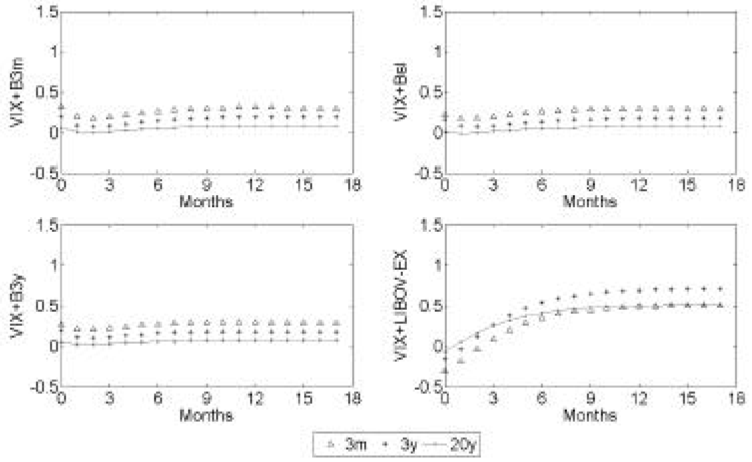


Figure 4. Impulse response of shocks to Fed factor on yields. This …gure shows the e¤ect of a shock to Fed factor θ_1^{US} on the Brazilian sovereign yields with maturities of 3 months, 3 years and 20 years up to 18 months after the shock. The shock is one standard deviation of a monthly variation of the Fed factor. The responses are evaluated considering the bilateral model.

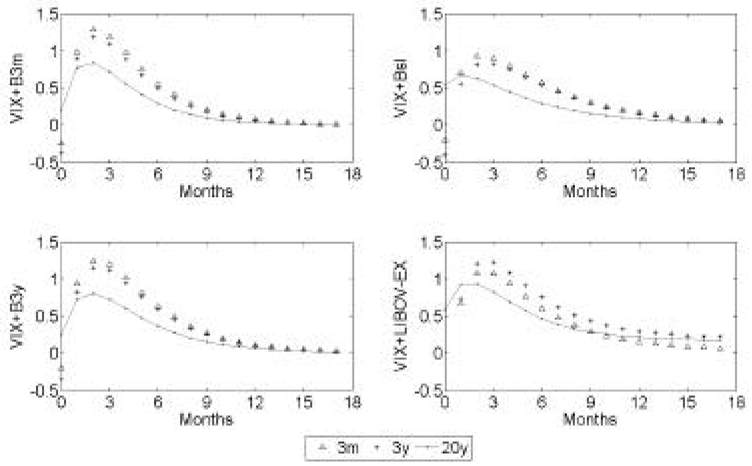


Figure 5. Impulse response of shocks to the VIX on yields. This …gure shows the e¤ect of a shock to the VIX on the Brazilian sovereign yields with maturities of 3 months, 3 years and 20 years up to 18 months after the shock. The shock is one standard deviation of a monthly variation of the VIX. The responses are evaluated considering the bilateral model.

allowing the use of important variables such as Central Bank reserves, real activity and in‡ation.

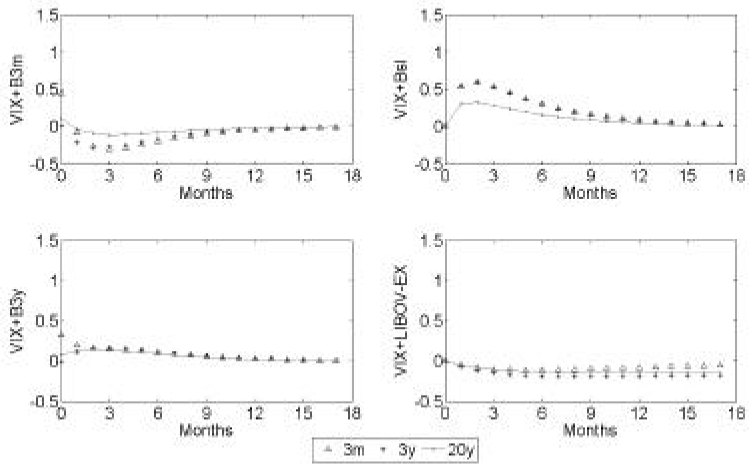


Figure 6. Impulse response of shocks to observable Brazilian factors on yields. This figure shows the effect of a shock to observable Brazilian factors on the Brazilian sovereign yields with maturities of 3 months, 3 years and 20 years up to 18 months after the shock. The shock is one standard deviation of a monthly variation of the observable factor. The responses are evaluated considering the bilateral model.

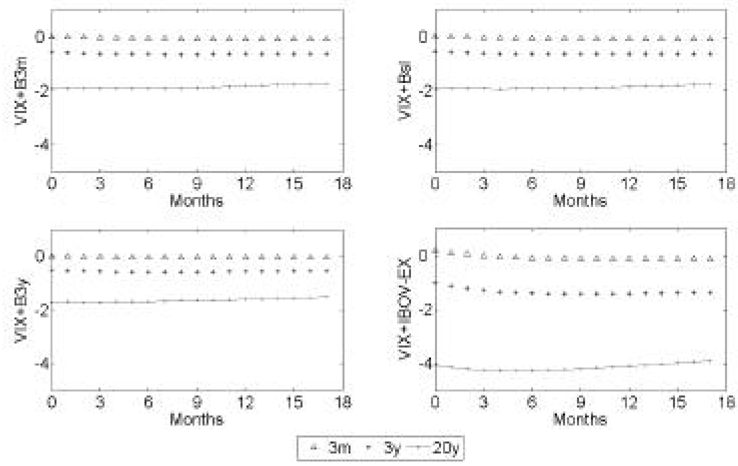


Figure 7. This figure shows the effect of a shock to Fed factor θ_1^{US} on the 3 months, 3 years and 20 years survival probabilities up to 18 months after the shock. The shock is one standard deviation of a monthly variation of the Fed factor. The responses are evaluated considering the bilateral model.

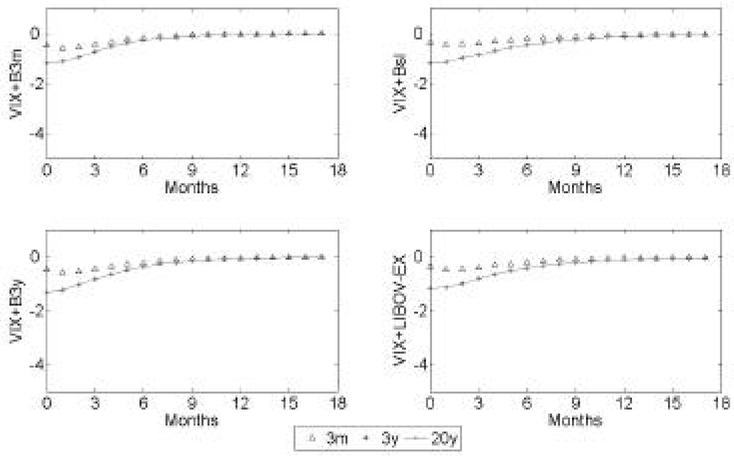


Figure 8. Impulse response of shocks to VIX on survival probabilities. This …gure shows the e¤ect of a shock to VIX on the 3 months, 3 years and 20 years survival probabilities up to 18 months after the shock. The shock is one standard deviation of a monthly variation of the VIX. The responses are evaluated considering the bilateral model.

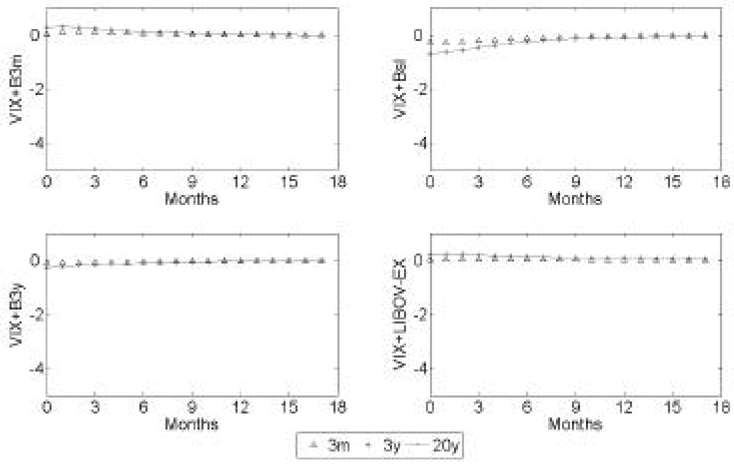


Figure 9. Impulse response of shocks to observable Brazilian factors on survival probabilities. This …gure shows the e¤ect of a shock to observable Brazilian factors on the 3 months, 3 years and 20 years survival probabilities up to 18 months after the shock. The shock is one standard deviation of a monthly variation of the observable factor. The responses are evaluated considering the bilateral model.

CHAPTER 3

Identi…cation of Term Structure Models with Observed Factors

1. Introduction

Multifactor term structure models possess better …tting and forecasting performance than single factor models, at the cost of demanding more computational time and greater samples for estimation. Hence, it is not surprising that many authors restrict the parameter space, sometimes using some economic criteria.

However, restrictions are actually necessary to identify the model: Dai and Singleton (2000), DS, showed that there are transformations of the parameters and state variables of the a¢ne models that preserve the yield curve. Therefore, multiple sub-identi…ed models correspond to the same data.

Arbitrary restrictions, on the other hand, may over-identify some parameters, distorting model properties. Thus, one must …nd a well-de…ned set of restrictions leading to exact identi…cations. There are 2 approaches in the a¢ne literature, the one by DS and another proposed by Du¢e and Kan (1996) and re…ned by Collin-Dufresne et al. (2006), which uses speci…c transformations of the state factors into observed state variables, thereby eliminating unnecessary parameters.

Our contribution relies on the observation that when observed factors are included in the state vector, new restrictions are needed that take into account the additional degrees of freedom given by the choice of the relative weight of the observed factors. Also, we extend the method to other term structure models: the common factor, the Nelson-Siegel, the Legendre and a proposed "quasi" a¢ne model.

We de…ne invariant operators for each model, and prove that they preserve the likelihood. Then, we show how to identify.

There are alternative identi…cation choices, which could in principle induce di¤erent model properties. However, we prove that the yield curve response is the same for any identi…cation choice. On the other hand, the latent factors response does depend on the choice.

Pericoli and Taboga (2006) also point out to one correctly identi…ed speci…cation of an a¢ne model with observed factors, but they require additional assumptions.

2. Models

Let $Y_t = (Y_t^1, .., Y_t^n)'$ be the vector of bond yields, $X_t = (M_t, \theta_t)'$ the state vector. $M_t \in \mathbb{R}^p$ and $\theta_t \in \mathbb{R}^q$ denote the observed and latent factors, respectively.

The general equations,

$$
\begin{aligned}
(2.1)\qquad Y_t^n &= A_n(.) + B_n^\theta(.)\theta_t + B_n^M M_t + \sigma_n u_t^n, \\
M_t &= \mu_M + \Phi_{MM} M_{t-h} + \Phi_{M\theta}\theta_{t-h} + \Sigma_{MM}\epsilon_t^M, \\
\theta_t &= \mu_\theta + \Phi_{\theta M} M_{t-h} + \Phi_{\theta\theta}\theta_{t-h} + \Sigma_{\theta M}\epsilon_t^M + \Sigma_{\theta\theta}\epsilon_t^\theta,
\end{aligned}
$$

accommodate the following 5 models:

- Affine (na): the factor weights are given by

  $$A_n = -a_n/n \text{ and } B_n = -b_n'/n,$$

  where a_n and b_n are obtained by recursive equations:

  $$
  \begin{aligned}
  a_1 &= -\delta_0,\ b_1 = -\delta_1, \text{ and, for every } n, \\
  b_{n+1}' &= -\delta_1' + b_n'(\Phi - \Sigma\lambda_1), \\
  a_{n+1} &= -\delta_0 + a_n + b_n'(\mu - \Sigma\lambda_0) + \frac{1}{2} b_n'\Sigma\Sigma' b_n.
  \end{aligned}
  $$

  We denote $\Phi^* = \Phi - \Sigma\lambda_1$ and $\mu^* = \mu - \Sigma\lambda_0$, the dynamics under Q.
- "Unrestricted" affine or "quasi" affine (qn). It is more flexible and easier to estimate than the affine model, but relaxes no-arbitrage conditions. B_n is calculated as above, but A_n is now determined such that

  $$\sigma u_t = Y_t - A - B_\theta\theta_t - B_M M_t$$

  has zero mean.
- Common factor (cf): unrestricted A, B.
- Nelson-Siegel (ns): A and B^M are unrestricted, and

$$(2.2)\qquad B_n^\theta = \left(1, \left(1 - e^{-\gamma n}\right)/\gamma n, \left(1 - e^{-\gamma n}\right)/\gamma n - e^{-\gamma n}\right).$$

- Legendre (lg): A and B^M are unrestricted, and

  $$B_n^\theta = \left(1, x, (3x^2 - 1)/2, (5x^2 - 3x)/2\right),$$

  where $x = 2n/N - 1$ and N is the longest maturity.

The set of parameters in each case is:

$$
\begin{aligned}
\text{na} &: \Theta = (\delta_0, \delta_1, \lambda_0, \lambda_1, \mu, \Phi, \Sigma, \sigma); \\
\text{qn} &: \Theta = (\delta_1, \lambda_1, \mu, \Phi, \Sigma, \sigma); \\
\text{cf} &: \Theta = (A, B, \mu, \Phi, \Sigma, \sigma); \\
\text{ns} &: \Theta = (A, B^M, \gamma, \mu, \Phi, \Sigma, \sigma); \\
\text{lg} &: \Theta = (A, B^M, \mu, \Phi, \Sigma, \sigma).
\end{aligned}
$$

3. Identification

We begin with the definition of the invariant operators. Consider a non-singular matrix $L \in \mathbb{R}^{(p+q)\times(p+q)}$,

$$(3.1)\qquad L = \begin{pmatrix} I & 0 \\ \alpha & \beta \end{pmatrix},$$

where $I \in \mathbb{R}^{p \times p}$ is the identity, and $\alpha \in \mathbb{R}^{p \times q}$, $\beta \in \mathbb{R}^{q \times q}$. The operator T_L acts on $\{\Psi, X\}$:

$$
\begin{aligned}
\text{na} &: \{\Psi, X\} \mapsto \{(\delta_0, (L')^{-1}\delta_1, \lambda_0, \lambda_1 L^{-1}, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\};\\
\text{qn} &: \{\Psi, X\} \mapsto \{((L')^{-1}\delta_1, \lambda_1 L^{-1}, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\};\\
\text{cf} &: \{\Psi, X\} \mapsto \{(A, BL^{-1}, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\};\\
\text{ns} &: \{\Psi, X\} \mapsto \{(A, B^M - B^\theta \beta^{-1}\alpha, \gamma, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\};\\
\text{lg} &: \{\Psi, X\} \mapsto \{(A, B^M - B^\theta \beta^{-1}\alpha, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\}.
\end{aligned}
$$

In the ns and lg cases $\beta = I$, because B^θ is fixed.

T_L transforms the latent factor and adjusts its "macro" content:

$$\theta_t \mapsto \alpha M_t + \beta\theta_t. \tag{3.2}$$

Let $\nu \in \mathbb{R}^{p+q}$, $\nu = (0, \nu^\theta)'$. The operator T_ν shifts the latent factors:

$$
\begin{aligned}
\text{na} &: \{\Psi, X\} \mapsto \{(\delta_0 - \delta_1'\nu, \delta_1, \lambda_0 - \lambda_1\nu, \lambda_1, \mu + (I - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\};\\
\text{qn} &: \{\Psi, X\} \mapsto \{(\delta_1, \lambda_1, \mu + (I - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\};\\
\text{cf} &: \{\Psi, X\} \mapsto \{(A - B\nu, B, \mu + (I - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\};\\
\text{ns} &: \{\Psi, X\} \mapsto \{(A - B\nu, B^M, \gamma, \mu + (I - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\};\\
\text{lg} &: \{\Psi, X\} \mapsto \{(A - B\nu, B^M, \mu + (I - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\}.
\end{aligned}
$$

The last operator, T_O, where $O \in \mathbb{R}^{d \times d}$ is a rotation matrix, rotates the parameters related to the Brownian motion and preserves X:

$$\Sigma \mapsto \Sigma O'. \tag{3.3}$$

T_O is related to the VAR ordering of the state variables with respect to endogeneity.

The operators are constructed so as to preserve the short rate, the risk premium and the observed factors. In fact, they preserve the likelihood, as is shown next. The format of the likelihood depends on the type of treatment for the stochastic singularity that appears when there are more maturities than latent factors. The case with Chen-Scott inversion was discussed in chapter 2.

Proposition 2. Given initial conditions $\tilde{m}_0 = Lm_0$ and $\tilde{C}_0 = LC_0L'$, the invariant transformations preserve the likelihood with the Kalman Filter.

Proof. We show that the marginal and full-information likelihoods are preserved, i.e.,

$$L(Y|M, \Psi) = L(Y|M, \tilde{\Psi}) \text{ and } L(Y|X, \Psi) = L(Y|\tilde{X}, \tilde{\Psi}),$$

where $(\tilde{X}, \tilde{\Psi})$ are the transformed variables and parameters. This is easily seen for the full information likelihood, which is given by

$$L(Y|X, \Psi) = \prod_t f(Y_t|\theta_t, M_t, \Psi) \tag{3.4}$$

$$= \prod_t \exp\left\{-1/2\left[\ln|\sigma\sigma'| + (Y_t - A - BX_t)'(\sigma\sigma')^{-1}(Y_t - A - BX_t)\right]\right\}. \tag{3.5}$$

We turn to the marginal likelihood case, given by

$$L(Y|M, \Psi) = \prod_t f(Y_t|Y_{t-1}, M_{t-1}, \Psi)$$

$$= \prod_t \exp\left\{-1/2\left[|Q_t| + (Y_t - f_t)'Q_t^{-1}(Y_t - f_t)\right]\right\}, \tag{3.6}$$

which depends on the Kalman filter equations (see West and Harrison, 1997): Given $X_0 \sim N(m_0, C_0)$, the prior of the state variables is defined as

$$X_t|Y_{t-1}, M_{t-1}, \Psi \sim N(a_t, R_t),$$

the forecast of the yields as

$$Y_t|Y_{t-1}, M_{t-1}, \Psi \sim N(f_t, Q_t),$$

and the posterior of the variables as

$$X_t|Y_t, M_t, \Psi \sim N(m_t, C_t),$$

where

$$\begin{aligned} a_t &= \mu + \Phi m_{t-1}, \\ R_t &= \Phi C'_{t-1}\Phi' + V, V = \Sigma\Sigma', \\ f_t &= A + Ba_t, \\ Q_t &= BR_tB' + \sigma'\sigma, \end{aligned}$$

where $V = \Sigma\Sigma'$, and A and B are the factor weights. Also, due to the fact that M_t is observable,

$$m_t = (M_t, m_t^\theta) \text{ and } C_t = \begin{pmatrix} 0 & 0 \\ 0 & c_t^\theta \end{pmatrix},$$

where

$$\begin{aligned} m_t^\theta &= a_t^\theta + [R_tB'Q_t^{-1}(Y_t - f_t)]^\theta, \\ c_t^\theta &= [R_t + R_tB'Q_t^{-1}BR_t']^{\theta\theta}. \end{aligned}$$

Now, consider the operator T_L. Observe that $\tilde V = \tilde\Sigma\tilde\Sigma' = LVL'$, $\tilde\mu^\star = L\mu^\star = L\mu + L\Sigma\lambda_0 = \tilde\mu + \tilde\Sigma\lambda_0$, $\tilde\Phi^\star = L\Phi^\star L^{-1} = L(\Phi + \Sigma\lambda_1)L^{-1} = \tilde\Phi + \tilde\Sigma\tilde\lambda_1$. We prove by induction that if it is assumed that $\tilde m_0 = Lm_0$ and $\tilde C_0 = LC_0L'$, then

(3.7) $\qquad \tilde m_t = Lm_t,\ \tilde C_t = LC_tL',\ \tilde a_t = La_t,\ \tilde R_t = LR_tL',\ \tilde f_t = f_t,\ \tilde Q_t = Q_t.$

For $t = 1$, we have

$$\begin{aligned} \tilde a_1 &= \tilde\mu + \tilde\Phi\tilde m_0 = L\mu + L\Phi L^{-1}Lm_0 = La_1, \\ \tilde R_1 &= \tilde\Phi\tilde C_0\tilde\Phi' + \tilde V = L\Phi L^{-1}LC_0L'L'^{-1}\Phi L' + LVL^{-1} = LR_1L', \\ \tilde f_1 &= \tilde A + \tilde B\tilde a_1 = A + BL^{-1}La_1 = f_1, \\ \tilde Q_1 &= \tilde B\tilde R_1\tilde B' + \tilde\sigma = BL^{-1}LR_1L'(L^{-1})'B' + \sigma = BR_1B' + \sigma = Q_1, \\ \tilde m_1^\theta &= [\tilde a_1 + \tilde R_1\tilde B'\tilde Q_1^{-1}(Y_1 - \tilde f_1)]^\theta = [La_1 + LR_1L'(L^{-1})'B'Q_1^{-1}(Y_1 - f_1)]^\theta, \\ \tilde c_1^\theta &= [\tilde R_1 + \tilde R_1\tilde B\tilde Q_1^{-1}\tilde B\tilde R_1']^{\theta\theta} = [LR_1L' + LR_1L'(L^{-1})'B'Q_1^{-1}BL^{-1}LR_1'L']^{\theta\theta}. \end{aligned}$$

Hence $\tilde m_1^\theta = [Lm_1]^\theta$ and $\tilde c_1^\theta = [LC_1L']^{\theta\theta}$, and therefore $\tilde m_1 = Lm_t$ and $\tilde C_t = LC_tL'$. Thus, for $t = 1$ (3.7) holds. Now, suppose property (3.7) is true for t. Then, analogous calculations show that (3.7) holds for $t + 1$. It follows that $L(Y_t|M, \Psi) = L(Y_t|M, \tilde\Psi)$.

Similar calculations prove that T_ν and T_O preserve $L(Y|M, \Psi)$. ∎

Thus, we proved that the operators indeed represent the degeneracy of the likelihood. The next result will show how to identify the models, by eliminating the degrees of freedom α, β, ν, O of the operators.

Proposition 3. The following procedure identifies the models.

Proof. i) For all models, the operator T_ν can be used to set $\mu^\theta = 0$, and the operator T_O to impose a lower triangular form for Σ.

ii) However, the use of T_L will depend on the type of model. Consider the na case. The identification is achieved by spending the restrictions on any parameter, but mainly on Σ, Φ, $\Phi^\star$ or λ_1.

iii) Restriction on Σ: α and β are chosen such that $\tilde{\Sigma}_{\theta M} = 0$ and $\tilde{\Sigma}_{\theta\theta} = \mathrm{I}$, i.e.,

$$L\Sigma = \begin{pmatrix} \mathrm{I} & 0 \\ \alpha & \beta \end{pmatrix} \begin{pmatrix} \Sigma_{MM} & 0 \\ \Sigma_{\theta M} & \Sigma_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \Sigma_{MM} & 0 \\ 0 & \mathrm{I} \end{pmatrix},$$

where $\tilde{\Sigma} = L\Sigma$. This does not exhausts the free parameters. So, we apply the transformation $T_S := T_L|\alpha = 0, \beta = S$ rotation matrix, such that $\Phi_{\theta\theta}$ or $\Phi^\star_{\theta\theta}$ becomes lower triangular. The operator T_S commutes with $L\Sigma$ so that it will only rotate orthogonal Brownian motions, preserving other parameters. This completes the identification. Another option is obtained by choosing β such that $\Sigma_{\theta\theta}$ is diagonal and then apply $T_D := T_L|\alpha = 0, \beta = D$ diagonal matrix, such that $\delta_\theta = 1$:

$$(D^{-1})^{\mathrm{I}}\delta = \begin{pmatrix} I & 0 \\ 0 & D^{-1} \end{pmatrix} \begin{pmatrix} \delta_M \\ \delta_\theta \end{pmatrix} = \begin{pmatrix} \delta_M \\ 1 \end{pmatrix}.$$

iv) Restriction on λ_1:

$$\lambda_1 L^{-1} = \begin{pmatrix} \lambda_1^{MM} & \lambda_1^{M\theta} \\ \lambda_1^{\theta M} & \lambda_1^{\theta\theta} \end{pmatrix} \begin{pmatrix} \mathrm{I} & 0 \\ -\beta^{-1}\alpha & \beta^{-1} \end{pmatrix} = \begin{pmatrix} \tilde{\lambda}_1^{MM} & \tilde{\lambda}_1^{M\theta} \\ 0 & \mathrm{I} \end{pmatrix},$$

so that we eliminate $\lambda_1^{\theta M}$ and $\lambda_1^{\theta\theta}$. Similarly, we could restrict λ_0 instead of μ.

v) Restriction on Φ or $\Phi^\star$. This case is harder, because it involves quadratic equations. To avoid this, we assume that Φ can be decomposed as $P\Lambda P^{-1}$, where Λ is a real diagonal matrix. If Φ has real and distinct eigenvalues, this is always possible. Since $\Phi = P\Lambda P^{-1}$, then $L\Phi L^{-1} = LP\Lambda(LP)^{-1}$. Now, α can be chosen such that LP becomes upper block triangular (ubt), that is,

$$LP = \begin{pmatrix} \mathrm{I} & 0 \\ \alpha & \beta \end{pmatrix} \begin{pmatrix} P_{MM} & P_{M\theta} \\ P_{\theta M} & P_{\theta\theta} \end{pmatrix} = \begin{pmatrix} P_{MM} & P_{M\theta} \\ 0 & P_{\theta\theta} \end{pmatrix}.$$

Then $(LP)^{-1}$ and the transformed $\tilde{\Phi} = L\Phi L^{-1}$ will also be ubt. That is, $\tilde{\Phi}_{\theta M} = 0$. The first identification of this type is the following. Choose β such that $\beta P_{\theta\theta}$ is a diagonal matrix, then $\tilde{\Phi}_{\theta\theta}$ will also be diagonal. Next, use T_D such that $\delta_\theta = 1$, completing the identification. Alternatively, we could have started with $\Phi^\star = P\Lambda^\star P^{-1}$. Yet another possibility is to impose a lower triangular $\beta P_{\theta\theta}$, which implies in a lower triangular $\tilde{\Phi}_{\theta\theta}$. Then use $T_T := T_L|\alpha = 0, \beta = T$ lower triangular matrix to impose a diagonal $\Sigma_{\theta\theta}$. Finally, the identification finishes by applying T_D such that $\delta_\theta = 1$.

vi) For the other models, the procedure is entirely analogous, except that in the case of lg and ns models, there are less degrees of freedom, for $\beta = I$. ¤

Table 1 shows a summary of identifications for the affine model. Type 2 assume that Φ or $\Phi^\star$ can be decomposed as $P\Lambda P^{-1}$. Otherwise, we call Type 1.

Table 1. Summary of identifications (affine case).

	$\Phi_{\theta M}$	$\Phi_{\theta\theta}$	$\Phi^{\star}_{\theta M}$	$\Phi^{\star}_{\theta\theta}$	$\Sigma_{\theta M}$	$\Sigma_{\theta\theta}$	δ^{θ}	$\lambda^{1}_{\theta M}$	$\lambda^{1}_{\theta\theta}$
Type1	Full	Full	Full	L Tr	0	I	Full	-	-
Type1	Full	Full	Full	L Tr	0	Diag	I	-	-
Type1	Full	L Tr	Full	Full	0	I	Full	-	-
Type1	Full	L Tr	Full	Full	0	Diag	I	-	-
Type1	Full	Full	-	-	Full	L Tr	Full	0	I
Type1	-	-	Full	Full	Full	L Tr	Full	0	I
Type2	Full	Full	0	Diag	Full	L Tr	I	-	-
Type2	Full	Full	0	L Tr	Full	Diag	I	-	-
Type2	0	Diag	Full	Full	Full	L Tr	I	-	-
Type2	0	L Tr	Full	Full	Full	Diag	I	-	-

Since there are many options, it is not clear how to choose a specification. At least, our next proposition shows that the response of the yield and observed variables to state factor shocks (IRFY, IRFM) and the variance decomposition of the yield (VDY) is the same for any case. In practice, however, at the inference stage, some alternatives may prove easier to be estimated.

Proposition 4. T_L and T_ν preserve IRFY, VDY and IRFM. All operators preserve the yield curve.

Proof. First note that for the qn, cf, ns and lg models, the invariance of the yields is immediate. Consider the na model. The weights $A, B|\{$ª$, X\}$ change to $A, BL^{-1}|T_L\{$ª$, X\}$ when T_L is applied, for $b'_{n+1}L^{-1} = -\delta'_1 L^{-1} + b'_n L^{-1}(L\Phi L^{-1} - L\Sigma\lambda_1 L^{-1})$ and $a_{n+1} = -\delta_0 + a_n + b'_n L^{-1}(L\mu - L\Sigma\lambda_0) + \frac{1}{2} b'_n L^{-1} L\Sigma\Sigma' L' (L^{-1})' b_n$. It follows that $Y_t = A + BL^{-1}LX_t + \sigma u_t$, which proves the invariance under T_L.

Furthermore, similar computations show that $A, B|\{$ª$, X\}$ changes to $A - B\nu, B|T_\nu\{$ª$, X\}$. Thus, $a_{n+1}|\{$ª$, X\}$ becomes $a_{n+1} - b'_{n+1}\nu|T_\nu\{$ª$, X\}$, so that $Y_t = A - B\nu + B(X_t + \nu) + \sigma u_t$, which proves the invariance under T_ν.

Now, IRFY, VDY$_j$ and IRFM, given by $\{B\Phi^n\Sigma\varepsilon_t\}_n$, $\{B\Sigma_j(B\Sigma_j)' + \ldots + B\Phi^n\Sigma_j(B\Phi^n\Sigma_j)'\}_n$ and the first p lines of $LB\Phi^n\Sigma\varepsilon_t$, are all easily seen to be preserved under T_L and T_ν. □

On the other hand, the response of the latent factors to factor shocks, IRFθ, depends on the identification: $L\Phi^n\Sigma\varepsilon_t$. This has implications on the interpretation of the latent factors, specially regarding monetary policy rules, since different specifications will give different "weights" to the macro content of the latent factors.

4. Conclusion

We specify invariant operators and prove that they preserve the likelihood of term structure models with observed factors. Hence, restriction choices must be made in order to identify the model, and we show ways of achieving it.

It turns out that any specification, if correctly estimated, lead to the same observable model properties such as the response of the yield curve to state variable shocks. However, some specifications can be numerically simpler to estimate. Also, the interpretation of the latent factors varies with the specification.

CHAPTER 4

Forecasting the Yield Curve: Comparing Models

1. Introduction

The modeling of the term structure of interest rates is a challenging task that has, from a practical perspective, at least three purposes. Firstly, with this tool available, one can price fixed income instruments and manage risk of bonds and derivatives. Secondly, it allows monitoring of observed and unobserved economic variables such as risk premium, default risk, inflation and real activity. Finally, it is possible to forecast futures interest rates.

In this study, we address the latter issue using a rich class of linear factor models. Examples of users of yield curve forecasts are numerous. Treasuries manage the emission and maintenance of the stock of public debt, which continuously demands an assessment of current and future interest rates. Investors must track their portfolios performance against the cost of opportunity of investing in low risk bonds. Central Banks react to the expected inflation and economic activity adjusting the short rate, thereby affecting the whole curve.

Term structure models can be classified in different ways. If restrictions to the evolution of the yields are imposed in order to avoid risk-free profit opportunities, then the model is known as arbitrage-free. Otherwise the model is said to be purely statistical. Arbitrage-free models contain some ingredients arising from equilibrium models having thus a strong economic appeal. Seminal works within this class are Vasicek (1977), Cox et al. (1985) and Heath et al. (1992) while Nelson and Siegel (1987) and Svensson (1994) are pioneer works in the class of statistical models. Moreover, term structure models may directly include or not macroeconomic and financial factors driving the yield curve. Among others, we can cite the works of Ang and Piazzesi (2003), Diebold et al. (2005), Hördahl et al. (2006), and Ludvigson and Ng (2007) that use macroeconomic variables to model the term structure of interest rate. Finally, the relation between factors and interest rates may be linear or assume a more general specification. Examples of linear models are the class of affine models studied by Duffie and Kan (1996)[1], while Longstaff (1989) and Leippold and Wu (2003) constitute examples of non-linear models.

Although several works from macroeconomics, finance and econometrics have been devoted to term structure models, few of them analyzed the out-of-sample forecasting performance. Predictability questions regarding yield curve are first studied by Fama and Bliss (1987) who investigated the relationship between forward and future spot rates. More recently, Duffee (2002) documents that the affine models produced poor U.S. yields forecasts. Diebold and Li (2006) propose a two-stage model based on the Nelson and Siegel (1987) framework to forecast the U.S. term

[1]An affine model is an arbitrage-free term structure model, such that the state process X is an affine diffusion, and the yields are also affine in X.

structure that presented better results than some competing models. Nevertheless, Almeida and Vicente (2008) show that the inclusion of no-arbitrage conditions in a latent model improves the out-of-sample …t. Ang and Piazzesi (2003) …nd that an a¢ne model with macroeconomic variables outperforms the unrestricted VAR model containing the same observable factors. They also show that models with macro factors forecast better than models with only unobservable factors. In the same line, Hördahl et al. (2006) con…rm that the forecasting performance of a model with observable factors is superior to models based on latent factors.

The papers cited in the last paragraph deal with models where the interest rates are linear functions of state factors (observable or latent). Due to variety of reasons (the more important is the ease of implementation), linear models are nowadays the workhorse of yield curve modeling. However, to the best of our knowledge there is no study based on the same dataset that provides a full comparison of the in-sample and out-of-sample performance of di¤erent speci…cations of linear models. In this article we try to …ll this gap in the …nance literature. Although some recent studies have addressed the forecasting performance of di¤erent linear interest rate models[2], we believe that no one has implemented a comparative analysis as comprehensive as ours. We analyze arbitrage-free and purely statistical models, with or without observable variables. In addition to testing the models with U.S. data as usual, we also consider a database from a emerging country. For instance, Laurini and Hotta (2007), Pinheiro et al. (2007), Almeida et al. (2007A, 2007B) and Shoucha (2005) present models estimated for Brazilian data.

All the models analyzed in this study present constant volatility. Although stochastic volatility processes have some desirable properties, the standard approach in the interest rates forecasting has been using homocedastic models. Besides the parsimonious, constant volatility models seems to be a natural choice when the aim is forecasting since in this family there is no factor collecting information about the volatility process. Therefore, it is expected that the mean of the yields distribution can be better captured.

We estimated the models via MCMC (see chapter 1), which will allow the measurement of the uncertainty associated to the information available for the estimation of a given model.

2. Models

As before, Y denote the vector of n interest rate yields, $X = (M, \theta)$ the state vector composed of p macro factors and q latent factors. As explained in the previous chapter, the general equations for the latent factor models are:

$$Y_t = A(.) + B^{\theta}(.)\theta_t + B^M M_t + \sigma u_t, \; u_t \sim N(0, I_n), \tag{2.1}$$

$$X_t = \mu + \Phi X_{t-h} + \Sigma \varepsilon_t, \; \varepsilon_t \sim N(0, I_{p+q}), \tag{2.2}$$

where h denotes the lag size, and $A(.)$ and $B(.)$ di¤erentiate the models.

For the Fama-Bliss case, the data is monthly and $h = 1$. The other samples are daily and $h = 1, 5$ or 21. This last choice turns a daily model into a sum of monthly models: since there are on average 21 business days in a month, when $h = 21$, the sample can be decomposed into 21 disjoint series and the full data

[2]Besides the works cited above, we can also include Vicente and Tabak (2008) and Moench (2008).

loglikelihood L can be thought of as the sum of 21 monthly loglikelihoods. Denote by the superscript d the monthly model estimated with data formed with the day d of each month:

$$L(\Psi, \theta | M, Y) = \sum_{d=1}^{21} L^d(\Psi^d, \theta^d | M^d, Y^d). \tag{2.3}$$

It is assumed above that the monthly models are independent, but that the set of parameters is the same: $\Psi^d = \Psi$. In this way, we estimate a monthly model using all available data, and not a particular sub-sample or average. Likewise, DL use monthly data and 6 or 12-month lags.

Next, we discuss the versions of the Nelson-Siegel (ns) and affine (na) models that are considered here.

2.1. $B^M = 0$. All the models are estimated fixing $B^M = 0$. This can be justified in the affine model as a choice of monetary policy, as explained in what follows.

In the affine model, the parameters are given by $A = [A_1; ..; A_N]$, $B = [B_1; ..; B_N]$, $A_n = -\alpha_n/n$ and $B_n = -\beta_n'/n$, where

$$\begin{aligned} \beta_{n+1}' &= -\delta_1'(I + \Phi^\star + .. + \Phi^{\star n}), \\ \alpha_{n+1} &= -\delta_0 + \alpha_n + \beta_n'\mu^\star + \frac{1}{2}\beta_n'\Sigma\Sigma'\beta_n, \end{aligned} \tag{2.4}$$

with initial condition $\alpha_1 = -\delta_0$.

However, as we saw in chapter 3, some restrictions are necessary to identify the models, but there are many equivalent alternatives to do it. In particular $\Phi^\star_{\theta M} = 0$, $\Phi^\star_{\theta\theta} = 0$, $\delta_1^\theta = 1$, in which case $\Phi^\star$ and $(I + \Phi^\star + .. + \Phi^{\star n})$ are upper triangular.

It turns out that if we also impose $\delta_1^M = 0$, in which case the short rate follows an infinite no-discount forward looking Taylor rule (see Ang et al., 2007), then $B^M = 0$.

Hence, $B^M = 0$ is consistent with a Taylor rule choice of an identified affine model.

2.2. Unrestricted. We propose "unrestricted" versions of the affine and Nelson-Siegel models. In the case of the affine models, B is calculated as usual, but A is now unrestricted. Then, since $B^M = 0$ and we impose that the stochastic process θ has zero mean for identification purposes, the estimation can be simplified by fixing $A = \overline{Y}$, so that equations (2.1,2.2) are replaced by

$$Y_t = \overline{Y} + B^\theta(.)\theta_t + \sigma u_t,\ u_t \sim N(0, I_n),$$

$$X_t = \Phi X_{t-h} + \Sigma\varepsilon_t,\ \varepsilon_t \sim N(0, I_{p+q}).$$

By subtracting the mean of the observed indicators, we have $\mu^M = 0$ and consequently $\mu = 0$. This model is called unrestricted affine.

Now, consider the standard Nelson-Siegel model:

$$Y_t = B^\theta(\gamma)\theta_t + \sigma u_t,\ u_t \sim N(0, I_n), \tag{2.5}$$

$$X_t = \mu + \Phi X_{t-h} + \Sigma\varepsilon_t,\ \varepsilon_t \sim N(0, I_{p+q}). \tag{2.6}$$

In the unrestricted version, the parameter vector A is added, which is then set equal to $\overline{Y}$ as in the unrestricted affine model. In this case, to be identified, it is necessary to impose $\mu = 0$. The equations then become

$$Y_t = \overline{Y} + B^\theta(.)\theta_t + \sigma u_t,\ u_t \sim N(0, I_n), \tag{2.7}$$

$$X_t = \Phi X_{t-h} + \Sigma\varepsilon_t,\ \ \varepsilon_t \sim N(0, I_{p+q}). \tag{2.8}$$

Note that by adding A, we can exactly fit the long term mean of the yield curve and enforce that the measurement errors u_t have zero mean, which is not possible in the standard case with $A = 0$ and free μ.

2.3. **Pooled versus aggregated.** The pooled model combines daily yield series and monthly macro series. This is discussed in chapter 5.

2.4. **Other models.** Our list of models also include autoregressive models, which are estimated in univariate or multivariate versions, and yields-alone or macro-augmented versions:

$$\text{AR: } Y_t^n = \mu^n + \Phi^n Y_{t-h}^n + u_t^n, \tag{2.9}$$

$$\text{VAR: } Y_t^n = \mu^n + \Phi^n Z_{t-h} + u_t^n,\ Z_t = (Y_t^1, Y_t^{N/2}, Y_t^N). \tag{2.10}$$

N denotes the N-th longest maturity, $Y^{N/2}$ indicates the yield with the $\lfloor N/2 \rfloor$-th maturity. This model is in fact a simplified VAR. The full VAR would require a too high number of varibles and parameters. In the macro-augmented VAR, $Z_t = (Y_t^1, Y_t^{N/2}, Y_t^N, M_t)$ include macro factors.

Completing the list, there are the two-stage Diebold and Li (2006) type of models: first, one estimates θ_t for every t in

$$Y_t = B^\theta \theta_t + \sigma u_t,\ \ u_t \sim N(0, I_n). \tag{2.11}$$

B^θ is either given by Laguerre weights (Nelson-Siegel) or Legendre polynomials (see Almeida et al., 1998). Then, conditional on $\hat{\theta}_t$, the other parameters are estimated via a AR or VAR:

$$\hat{\theta}_t = \mu + \Phi\hat{\theta}_{t-h} + \Sigma\varepsilon_t,\ \ \varepsilon_t \sim N(0, I_{p+q}). \tag{2.12}$$

For the version with macro factors, we have instead:

$$\begin{aligned} M_t &= \mu^M + \Phi^{MM} M_{t-h} + \Phi^{M\theta}\hat{\theta}_{t-h} + \Sigma^{MM}\varepsilon_t^M, \\ \hat{\theta}_t &= \mu^\theta + \Phi^{\theta M} M_{t-h} + \Phi^{\theta\theta}\hat{\theta}_{t-h} + \Sigma^{\theta M}\varepsilon_t^M + \Sigma^{\theta\theta}\varepsilon_t^\theta. \end{aligned} \tag{2.13}$$

3. Inference

The latent factor models are estimated via Monte Carlo Markov Chain (MCMC), described in chapter 1 for the case of affine models. We now present the algorithm for the common factor (cf), the Nelson-Siegel (ns), the unrestricted ns (ns_u), and the unrestricted affine (affine_u) models.

We divide the set of parameters $\Theta = (\mu, \Phi, \sigma, \zeta)$ - where ζ depends on the type of model - into subsets that can be analytically sampled when possible. Otherwise, it is sampled with Metropolis-Hastings:

cf	ns/ns_u	affine	affine_u
$\zeta = (A, B, \Sigma)$	$\zeta = (\gamma, \Sigma)$	$\zeta = (\delta_0, \delta_1, \mu^\star, \Phi^\star, \Sigma)$	$\zeta = (\delta_1, \Phi^\star, \Sigma)$

Starting from an initial vector (Θ^0, θ^0), the algorithm consists of the repeated sampling of Θ given the observable data until some convergence criteria is met. In the end, a distribution of parameters is obtained.

For the k-th iteration, we have:

(1) Draw $(\mu^k, \Phi^k) \sim p(\mu, \Phi | \sigma^{k-1}, \zeta^{k-1}, \theta^{k-1}, Y, M)$,

(2) Draw $\sigma^k \sim p(\sigma|\mu^k, \Phi^k, \zeta^{k-1}, \theta^{k-1}, Y, M)$,
(3) Draw $\theta^k \sim p(\theta|\mu^k, \Phi^k, \sigma^k, \zeta^{k-1}, Y, M)$,
(4) Draw $\zeta_i^k \sim p(\zeta_i|\mu^k, \Phi^k, \sigma^k, \theta^k, \zeta_{-i}^{k-1}, Y, M)$.

Thus, the estimation problem is decomposed into a series of simpler sub-problems. Sub-problems 1-3 use Gibbs sampling, and correspond to, respectively, the estimation of a VAR model, the estimation of its variance, and the joint distribution of latent factors:

Subproblem 1:

$$f(\mu, \Phi|\sigma^{k-1}, \zeta^{k-1}, \theta^{k-1}, Y, M) \sim N((X'X)^{-1}X'X^\star, (X'X)^{-1} \otimes \Sigma), \tag{3.1}$$

where $X = (X_1, ..., X_{T-1})'$, $X^\star = (X_2, ..., X_T)'$, $X = (M, \theta)$.

Subproblem 2:

$$f(\sigma|\mu^k, \Phi^k, \zeta^{k-1}, \theta^{k-1}, Y, M) \sim IG(diag(U'U), n), \tag{3.2}$$

where $U = Y - A - BX$, and IG is the inverse gamma distribution.

Subproblem 3:

$$f(\theta|\mu^k, \Phi^k, \sigma^k, \zeta^{k-1}, Y, M). \tag{3.3}$$

This problem is solved via the FFBS algorithm (Carter and Kohn, 1994), seen in chapter 1.

The step relative to ζ is model specific:

Subproblem 4:

$$f(\zeta_i|\zeta_{-i}^{k-1}, \mu^k, \Phi^k, \sigma^k, \theta^k, Y, M). \tag{3.4}$$

- Σ: For the ns, affine_u and cf cases, there is a known distribution, given by the inverse wishart distribution. For the affine case, we sample using Independence Metropolis with an inverse wishart proposal: draw a candidate $\zeta_i^k \sim \zeta_i^{k-1} + N(0, c)$, where c is a constant. If

$$L(\zeta_i^k|\zeta_{-i}^{k-1}, \mu^k, \Phi^k, \sigma^k, \theta^k, Y, M) - L(\zeta_i^k|\zeta_{-i}^{k-1}, \mu^k, \Phi^k, \sigma^k, \theta^k, Y, M) > \log(z), \tag{3.5}$$

  where L is the log-likelihood and $z \sim U(0, 1)$, then accept it, otherwise $\zeta_i^k = \zeta_i^{k-1}$. By calibrating c, the acceptance ratio is maintained in the 20-50% range.
- $\delta_0, \mu^\star$: affine case. Usually, these parameters are estimated via Metropolis. However, as noticed by Johannes and Polson (2006), $\mu^\star$ appears in a linear way on the pricing equation. Based on this, we analytically derive a distribution for $\mu^\star$. Rearranging the recursive equations, one obtains:

$$\begin{aligned} B_n &= \delta_1'(I + \Phi^\star + .. + \Phi^{\star n-1})/n, \\ A_n &= \delta_0 + \eta_n\mu^\star + \zeta_n, \\ \eta_n &= (B_1 + .. + B_n), \\ \zeta_n &= -n(B_1\Sigma\Sigma'B_1' + .. + B_{n-1}\Sigma\Sigma'B_{n-1}'). \end{aligned} \tag{3.6}$$

  Hence

$$\begin{aligned} Y_t^n &= A_n + B_nX_t + \sigma_n u_t^n \\ &= \delta_0 + \eta_n\mu^\star + \zeta_n + B_nX_t + \sigma_n u_t^n, \\ \tilde{Y}_t^n &:= Y_t^n - B_nX_t - \zeta_n = \delta_0 + \eta_n\mu^\star + \sigma_n u_t^n, \end{aligned} \tag{3.7}$$

  so that $\delta_0, \mu^\star$ can be solved by generalized least squares.
- δ_1, $\Phi^\star$: affine case. Use random walk Metropolis with normal proposal.

- γ: ns case. Use random walk Metropolis with normal proposal.
- A: cf, affine_u and ns_u cases. Set $A = \overline{Y}$.
- B: cf case. The solution is a regression similar to that for μ, Φ.

3.1. Performance Criteria and specifications.

3.1.1. In-sample Information Criterion. Due to the fact that the models under investigation have a different number of parameters, it is not possible to compare absolute values of the likelihood. Therefore, we choose as our criteria two Bayesian measures that emphasize forecasting performance and adherence to data, taking into account the number of parameters. Gelfand and Ghosh (1998) proposed the minimum posterior predictive loss (PPL) criterion emphasizing forecasting performance and Spiegelhalter (2002) proposed the deviance of information criterion (DIC) emphasizing adherence.

- PPL: For each realization of the distribution of the parameter estimators $\Theta^w \sim (\Theta|Y)$, there corresponds a yield curve forecast $Y|\Theta^w$. Consider a loss function penalizing both the expected error $E(Y|\Theta^w) - Y$ and the variance of the forecasts $Y|\Theta^w - E(Y|\Theta^w)$. In our case, the target variable is multivariate, so that we take the mean of the expected losses calculated for each maturity. In other words, the criterion is:

(3.8)
$$PPL = \sum_n \sum_t (Y_t^n - E(Y_t^n|D_{t-h}))^2 + \frac{1}{2}\sum_n \sum_t \frac{1}{N_w}\sum_w (E(Y_t^n|\Theta^w) - E(Y_t^n|D_{t-h}))^2,$$

where N_w is the number of realizations.

- DIC: It consists of a generalization of the Akaike criterion (AIC), based on the distribution of the divergence $D(\Theta) = -2\log L(\Theta)$:

$$DIC = E_w(D(\Theta_w)) - p_d = 2E(D(\Theta)) - D(E_w(\Theta_w)), \tag{3.9}$$

where $p_d = E(D(\Theta)) - D(E(\Theta))$ measures the equivalent number of parameters in the model, $E(D(\Theta))$ is the mean of the divergences calculated with the estimators' posterior distribution, and $D(E(\Theta))$ is the divergence calculated at the mean point of the estimators' posterior distribution.

The choice of the criterion should be linked to the objective of the estimation. Banerjee et al. (2004) show that PPL and DIC evaluates the goodness of fitting of the models penalizing for the degree of complexity. DIC considers the likelihood on the space of the parameters and PPL on the posterior predictive space. Thus, when the main interest lies in forecasting, PPL is preferred, whereas when the capacity of the model to explain the data is more desirable, DIC should be used.

Besides those measures, we will calculate models in-sample RMSE normalized by the random walk (RW) RMSE.

3.1.2. Out-of-sample Criterion. To compare the out-of-sample results, we define 3 measures using the Theil-U, the ratio between the RMSE of the model forecasts and the RMSE of the RW forecasts. When the processes under study have high persistency, RW frequently adheres well to data, sometimes even more than sophisticated models.

$$TU(n) = \left(\sum_{t_{out}} (Y_t^n - \hat{Y}_{t|t-h}^n)^2 / \sum_{t_{out}} (Y_t^n - Y_{t-h}^n)^2 \right)^{\frac{1}{2}}.$$

Note that another metric could be used to normalize the RMSE instead of the RW, the long term average:

$$\mathrm{TU(n)} = \left(\sum_{t_{out}} (Y_t^n - \hat{Y}_{t|t-h}^n)^2 / \sum_{t_{out}} (Y_t^n - \overline{Y}^n)^2 \right)^{\frac{1}{2}}.$$

Since a high number of models are estimated, we do not report the TU for every maturity and model. Instead, we count the number of maturities such that $E(\mathrm{TU}) < 1$, which means that the model prediction errors were in mean lower than the RW's. This number is the first out-of-sample criterion and is denoted by t.

The other two criterion are refinements of t. They count the maturities such that:

- $E(\mathrm{TU}) + 1.65\sigma(\mathrm{TU}) < 1$, which amounts to a statistically significant t and is denoted by s. This value can be computed because MCMC provides sample distributions of functions of parameters;
- $E(\mathrm{TU}) < 1$ and the Diebold-Mariano (see Diebold and Li, 2006) criterion is satisfied, which we call criterion d.

4. Data

From DIxPre interest rate swap contracts traded in the Brazilian Futures Market (BM&F), we obtain constant maturity zero-coupon rates of the Brazilian domestic term structure.

For the U.S. term structure, we use two sources. One is the daily yield series provided by the Federal Reserve, free of coupon and with constant maturity. The other is the Fama-Bliss sample, constituted of monthly constant maturity zero-coupon rates, and used by other authors such as DL and Vicente and Tabak (2008).

Table 1A. Yield data description.

	Brazil (BM&F)	U.S. (FED)	U.S. (Fama-Bliss)
in-sample	Jan 99 - Mar 06	Jan 1987 - May 03	Jan 85 - Jan 92
out-of-sample	Apr 06 - Mar 07	Jun 2003 - Mar 08	Fev 92 - Dec 00
forecast. hor.	1m, 6m	1m, 12m	1m, 12m
maturity (m)	1,2,3,6,9,12,18,24,36	1,3,6,12,24,36,60,84,120	3,12,36,60,120

Table 1B. Observed data source.

	Brazil	U.S.
C.P.I.	IBGE	FED
Industrial Output Gap	Authors' construction from IBGE	FED
Bovespa Index	Ipeadata	-

Each model has a daily data (estimated as explained with a 21 days lag) and a monthly data version. In this way, we are able to evaluate if the effect of sample reduction when passing from daily to monthly versions affect the estimation of the American and Brazilian cases.

The estimation of macro-finance versions face a trade-off between having a smaller set of high frequency variables to form larger samples, or having a larger set of lower frequency variables, such as macroeconomic variables, to form smaller samples. In this case, most of the yield information is usually discarded.

This trade-off is specially important for emerging countries, which have relatively shorter and less accurate historical series as compared to developed economies data, besides having much more frequent changes of regime, which further reduce data availability.

For the Brazilian market, we evaluate the importance of observed factors for forecasting by comparing versions with and without them. Our results indicated that, among the tested factors, only the IBovespa (São Paulo Stock Exchange Index) provided significant advantage.

The Brazilian sample begins in January 1999, shortly after a forced currency devaluation, or speculative attack, which occurred during the transition from fixed to floating exchange rate regime. Furthermore, some months later, in July 1999, the Inflation Target regime was established.

Since the modifications of the exchange rate regime and monetary policy affect the domestic prices formation mechanism, we decide to commence our sample from this event. The data for estimation ends in March 2006, and the forecasting exercise uses a further year of data.

Also, even very recent events greatly affected the domestic and sovereign bonds. In September 2002, one month before a presidential election in Brazil, the EMBI Plus country risk peaked 2436 basis points above the U.S. rate. In comparison, the same index remained around 200 points in 2007.

The U.S. series starts in 1987. In August of the same year, Greenspan was confirmed as the Chairman of the Federal Reserve, two months before the Crash of the NY Stock Exchange.

Among relevant events during the sample period, we list the LTCM collapse shortly after the Russian default in August 1998, which threatened the credit markets, the Internet bubble, and the terrorist attack in 2001 that led the FED to continuously lower the short rate until it reached 1%.

The proportion of the variance explained by the first 3 principal components of the correlation matrix is given below, and suggests that the markets have at most 3 sources of independent stochastic variance.

Table 2.

	First	Second	Third
BR: DIxPre swap	90.5	9.0	0.5
U.S.: FED	95.9	3.6	0.5

5. Estimation

Table 3 compares the deviance of information criterion (dic) and the posterior predictive loss (ppl) of the yields-only and of the macro-finance models, in the aggregated and pooled cases. It shows that for the American yield curve, the addition of macro factors does not improve the in-sample information. Also, dic and ppl criteria favour ns specifications over affine specifications. The values are close to the common factor model, suggesting that ns restrictions are already sufficiently flexible.

For the Brazilian data, the unrestricted affine and ns models scored higher than the standard versions. The best model under the dic criterion is the affine, and under the ppl the ns. Moreover, the pooled model did not improve the ppl criterion with respect to the aggregated model. On the other hand, the addition of the macro factors or the IBovespa improved somewhat the in-sample fitting of the monthly and pooled models.

Regarding the out-of-sample results, the best results for the American yield curve come from the unrestricted ns, which exhibits predictive capacity for 12 months ahead horizon. Neither the macro-…nance nor the pooled versions improved the results.

Table 3a. U.S.: dic and ppl in-sample information criteria. Models: common factor, a¢ne and Nelson-Siegel, unrestricted a¢ne and unrestricted ns.

	models	yields-only		macro-…nance	
	inf. criteria	dic	ppl	dic	ppl
monthly	common f.	-14267	0.253		
	a¢ne	-12484	0.536	-10878	0.520
	a¢ne_u	-13027	0.283	-12921	0.417
	ns	-13557	0.338	-13561	0.344
	ns_u	-13920	0.255	-13922	0.252
pooled	a¢ne_u	-263330	0.345		
	ns	-282800	0.334	-282660	0.332
	ns_u	-292360	0.253	-292310	0.249

Table 3b. Brazil: dic and ppl in-sample information criteria.

	models	yields-only		macro-…nance	
	inf. criteria	dic	ppl	dic	ppl
monthly	common f.	-3962	1.518		
	a¢ne	-3109	2.349	-3130	2.282
	a¢ne_u	-3920	1.563	-3923	1.650
	ns	-3787	1.968	-3787	2.065
	ns_u	-3876	1.547	-3885	1.516
pooled	a¢ne_u	-80755	1.631	-80861	1.651
	ns	-78395	1.931	-78497	1.958
	ns_u	-80012	1.513	-80181	1.474
	ns(bov)			-78671	1.908
	ns_u(bov)			-80038	1.511

Table4a. U.S. case. Summary of forecasts: number of maturities such that the TU is less than 1, for one month ahead (t1) or 12 months ahead (t12). The total number of maturities is 9. If in addition Diebold-Mariano test is satis…ed, or if TU plus one standard deviation is less than 1, it is counted in statistics d and s, respectively.

model		yields-only						macro-…nance					
		t1	t12	d1	d12	s1	s12	t1	t12	d1	d12	s1	s12
other	ar	9	0	3	0			4	0	5	1		
2 step	dl	0	0	3	0			4	0	5	0		
month	cf	3	7	2	0	0	0						
	a¢ne	0	0	0	0	0	0	0	0	0	0	0	0
	a¢ne_u	1	4	0	0	0	1	0	0	0	0	0	0
	ns	0	4	0	4	0	1	0	6	0	5	0	4
	ns_u	6	9	3	8	1	7	6	9	3	7	0	7
pool	a¢ne_u	0	0	0	0	0	0						
	ns	0	4	0	4	0	1	0	6	0	5	0	4
	ns_u	7	9	5	8	1	6	5	9	4	8	0	7

Contrary to the U.S. case, for Brazil the models generally showed low predictive capacity. The exceptions are the ns model with Bovespa, the Diebold-Li and AR models, all for the 6 months horizon.

The inclusion of the IBovespa substantially improved the long-term predictions, but not short term results. Also, macro-augmented and pooled versions result are similar to the standard monthly yields-only results.

Table4b. Brazilian case. Summary of forecasts: number of maturities such that the TU is less than 1, for one month ahead (t1) or 6 months ahead (t6). The total number of maturities is 9. If in addition Diebold-Mariano test is satis…ed, or if TU plus one standard deviation is less than 1, it is counted in statistics d and s, respectively.

model		yields-only						macro-…nance					
		t1	t6	d1	d6	s1	s6	t1	t6	d1	d6	s1	s6
other	ar	0	0	0	0			0	5	0	2		
2 steps	dl	0	0	0	0			0	5	0	2		
monthly	cf	0	2	0	0	0	0						
	a¢ne	0	0	0	0	0	0	0	0	0	0	0	0
	a¢ne_u	1	2	0	0	0	0	0	0	0	0	0	0
	ns	0	0	0	0	0	0	0	0	0	0	0	0
	ns_u	1	2	0	0	0	0	2	2	0	0	0	0
pooled	a¢ne_u	0	0	0	0	0	0	0	0	0	0	0	0
	ns	0	0	0	0	0	0	0	0	0	0	0	0
	ns_u	0	0	0	0	0	0	0	0	0	0	0	0
	ns(bov)							0	8	0	8	0	5

5.1. TU. The graphs below exhibit the values of the Theil-U for the 9 yield maturities, {1,3,6,12,24,36,60,84,120}-month for U.S. and {1,2,3,6,9,12,18,24,36}-month for Brazil. We present the results for the ns and dl models, comparing their speci…cations.

Each model is represented by two curves: the lower curve is the TU minus one standard deviation and the upper curve the TU plus one deviation. The pooled and the aggregated models are compared for the one month and the 12 months ahead forecasts using the Brazilian and American yield curve.

Figure A contains the U.S. forecasts. In the …rst row, the Diebold-Li (dl) model is compared to the pooled yields-only (dnsy) or macro-…nance (dnsr) models. The lower row considers the aggregated version. In both bases, we see that for the one month horizon, dl and ns models are similar, but for the 12 months horizon the dl model is outperformed by the ns versions. It also can be seen that all models exhibit good predictive capacity.

Figure B treats the Brazilian case. In the …rst row, dl, the aggregated (nsy) and the pooled (dnsy) versions are compared for the one and 6 months ahead predictions. In all cases, they are speci…ed as yields-only.

The second row shows the comparison about the choice of observed factor, IBovespa or in‡ation and output gap. It is seen that the only model with good forecasts is the Nelson-Siegel with IBovespa for the 6-month horizon. The use of macro factors or pooled versions do not help to improve the low performance of the yields-only model.

Figure A. TU U.S.: Pooled x Aggregated.

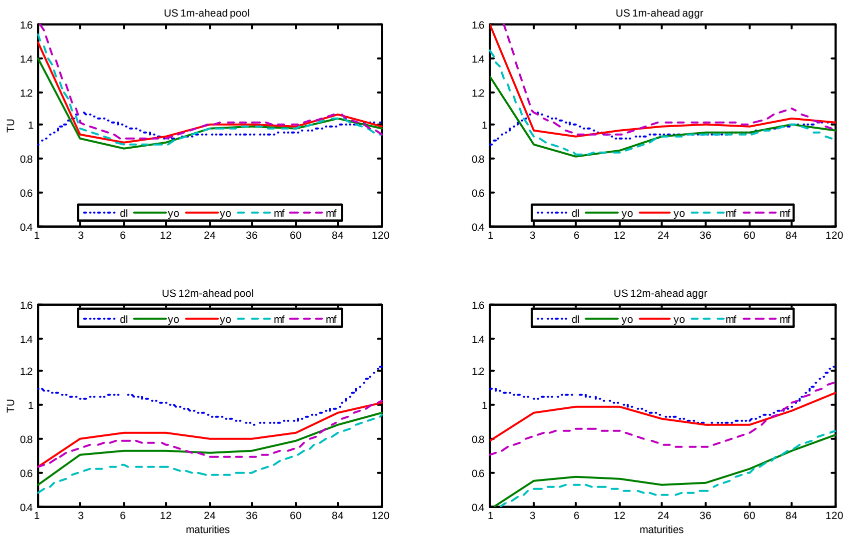


Figure 1

Figure B. TU BR: Macro-…nance x IBovespa.

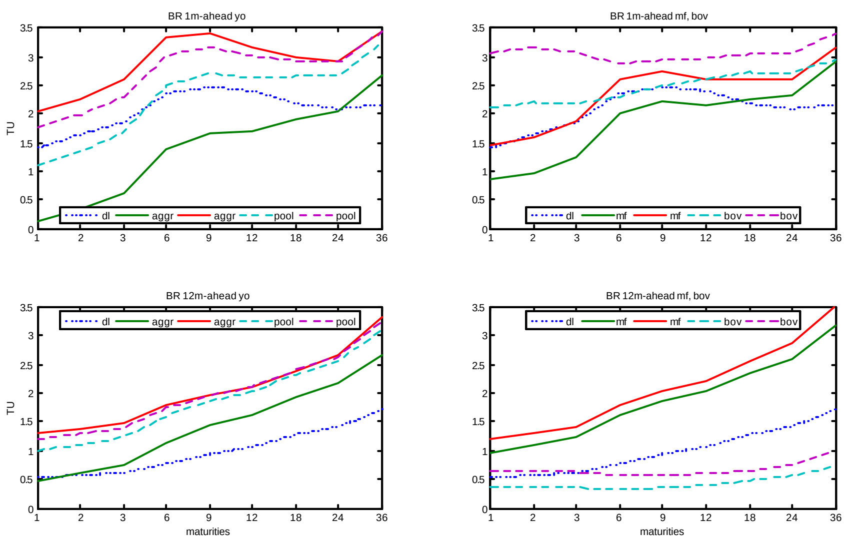
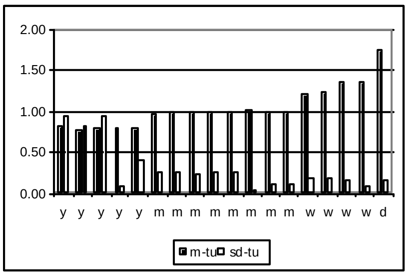
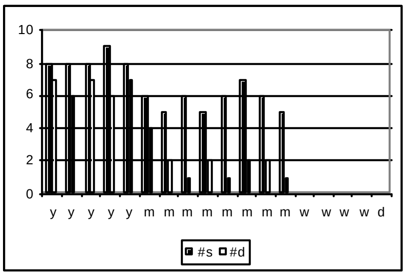


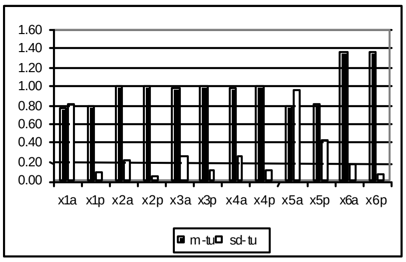
Figure 2

Next, we conduct an exercise in which we fixed the forecasting horizon and varied the size of the lag (in the graph, y denotes 1-year horizon, m 1-month, w 1-week and d 1-day). In the left graph below, we calculate average TU for the maturities (m-tu) and its standard deviation (sd-tu). The result is that for longer horizons we obtain good forecasts no matter what lag size is used, while short term predictions are generally poorer. In the right hand graphics, we compute the number of maturities that are statistically lower than 1 according to the MCMC sample (statistics s) or the Diebold-Mariano test for the same horizons (statistics d).





The graphics below compares the pooled and the aggregated models with respect to the average TU and its standard error. Here x1, x2, .., x6 denote different models, estimated in aggregated (a) or pooled (p) form. We see that the average TUs are not affected by the methodology, but as expected, the standard errors are always lower for the pooled versions.



Finally, we show below the results of the other models, composed of ar, var and the models estimated in 2 steps, namely, Nelson-Siegel and Legendre without or with macro, denoted dl and lg or dlv and lgv, respectively, and recursively estimated, denoted dlr and lgr. The forecasting performance of the two-step models were not higher than those of the jointly estimated latent factor models. As before, longer horizons results are superior to shorter horizons, particularly for var, nsv and lgv.

Table 5. Brazil. Two-step estimation. AR, Diebol-Li, Legendre.

monthly	ar	dl	lg	dlr	lgr	var	dlv	lgv
t1	0	0	0	0	1	0	0	0
t6	0	0	0	1	0	5	5	3

The last table presents results for the Fama-Bliss database. It shows the results of yields-only 2 and 3 latent factor models (common factor and Nelson-Siegel), VAR models, Diebold-Li model, ns and lg and macro-augmented versions nsv and lgv.

Table 6. Fama-Bliss: U.S. Latent factor, DL type models. 12 months ahead forecast.

Hor	cf2	ns2	cf3	ns3	ar	var	ns	lg	nsv	lgv	dl
3	0.77	1.05	0.94	1.35	0.96	1.31	1.39	0.71	0.85	0.75	0.73
12	0.81	0.89	0.86	1.33	0.96	1.58	1.23	0.74	0.72	0.74	0.70
36	0.79	0.94	0.79	1.37	0.80	1.73	1.46	0.75	0.83	0.80	0.74
60	0.89	1.11	0.87	1.56	0.80	1.89	1.67	0.82	0.97	0.89	0.82
120	1.20	1.47	1.17	1.97	1.41	2.46	2.06	0.93	1.16	0.98	0.93

6. Conclusion

Using MCMC, we estimated di¤erent classes of term structure models, in order to compare the out-of-sample forecasts using American and Brazilian data.

"Pooled" versions of the macro-…nance models that combine daily yield data and monthly macro data do not improve the forecasts, since their forecasts are equal on average to the standard models when there are su¢cient data, but the estimates and forecasts become more accurate.

The best model for the U.S. market was the unrestricted Nelson-Siegel. Either the yields-only or the macro-…nance versions were similar in this case.

The use of IBovespa improved the latent factor models for Brazil for 6 months ahead forecasts. DL models are easier to estimate, but the latent factor models, estimated in a more statistically sound way, showed at least as good results.

CHAPTER 5

Macro-Finance Models Combining Daily and Monthly Data

1. Introduction

The movements of the YC shape may reveal specific features about agents' expectations and influences from macroeconomic and financial indicators. After Ang and Piazzesi (2003) proposed a macro-finance model incorporating economic indicators in a term structure model, a number of articles followed, notably Ang et al. (2007), Rudebusch and Wu (2004), Hördahl et al. (2006), Dewachter and Lyrio (2006), Diebold et al. (2006) and others.

This article aims to contribute to that literature in methodological and empirical aspects. We propose a method that combines monthly and daily data. On the empirical side, we analyze the dynamic properties of the models by means of impulse response functions and variance decompositions, which are calculated for American and Brazilian markets.

We further propose an extension of the model that takes into account changes of regime on the reaction function of the yield curve to macroeconomic shocks. That is compared to an alternative way to accomplish the same task, by dividing the sample into sub-samples corresponding to the possible regimes and comparing the results. For example, in the Brazilian case, we test sub-samples corresponding to the president Cardoso after the adoption of the floating exchange rate and the inflation-target regime, and president Lula starting in 2003.

For the U.S. sample, sub-samples corresponding to Greenspan or Bernanke as chairman of the Federal Reserve are compared. However, in this case we show that the data in the Bernanke sub-sample is insufficient for the estimation of our models.

The models used here can be classified into two categories, the no-arbitrage and the econometric. In the first, the relation between the state variables and the YC is derived considering that it is not possible to arbitrage from rates of different maturities. This condition introduce restrictions that reduce the number of parameters at the cost of computationaly complicating the estimation.

In the econometric models, there are no such enforcements, but as Diebold and Li (2006) argues, if the data satisfies this condition and the model has good fitting, then this condition is indirectly met. In any case, the advantages and limitations of each approach, such as more formal rigor or clearer interpretation of the parameters, are empirically confronted here.

Once estimated, the models show the dynamic effect of the identified shocks on the YC. We compare the behavior of very different markets: U.S. possesses longer and more stable data, while Brazil, have shorter available data which may also exhibit multiple regimes.

In the macroeconomic literature, the forward versus backward looking approach is an unsolved issue in monetary policy theory (see Rudebusch and Wu, 2004). The macro-finance models, when combining macro variables and the YC, revisit this question. Since the YC is composed of variables driven by expectations, a natural way to include expectations into a dynamic macro model is to incorporate the YC.

As shown in the previous chapter, assuming an infinite no-discount forward-looking policy justifies for affine models a restriction that is equivalent to consider that the macro factors do not directly affect the yield curve in the observation equation.

Macro and financial variables have different temporal unities. The former are normally measured in monthly or quarterly frequency, which is the usual periodicity of the macro-finance models. Thus, the models ignore the almost continuous-time dynamics of the financial variables. Also, in the monthly models, the intra-month variability of the daily variables is lost – by taking the monthly mean or selecting the value observed for one day of the month –, with consequent loss of information.

We propose a "pooled" version of the macro-finance model that avoids this yield data, a feature that is particularly important for the case of developing markets which lack available samples.

The models were estimated with MCMC, with which we calculate information criteria that compare the performance of models with different degrees of complexity and number of parameters, and confidence intervals of the impulse response functions and variance decompositions.

Before the estimation of the models, we performed a simulation exercise, presented in the Appendix. We tested the MCMC algorithm using the simulated data and were able to recover the true parameters. Also, we calculate the Gelman-Rubin convergence diagnostics of the Markov chains generated by the Gibbs sampling and Metropolis-Hastings algorithms.

2. Term Structure

Let us rewrite the general term structure equations

$$Y_t = A + B_M M_t + B_\theta \theta_t + \sigma u_t,\ u_t \sim N(0, I_n), \tag{2.1}$$
$$M_t = \mu_M + \Phi_{MM} M_{t-h} + \Phi_{M\theta}\theta_{t-h} + \Sigma_{MM}\epsilon_t^M,\ \epsilon_t^M \sim N(0, I_p),$$
$$\theta_t = \mu_\theta + \Phi_{\theta M} M_{t-h} + \Phi_{\theta\theta}\theta_{t-h} + \Sigma_{\theta M}\epsilon_t^M + \Sigma_{\theta\theta}\epsilon_t^\theta,\ \epsilon_t^\theta \sim N(0, I_q),$$

substituting the time index t by the indices day d and month m.

Consider two alternatives for the observed factor: either it is collected with monthly frequency, in which case equation (2.3) is used, or it is collected on a daily basis, which corresponds to equation (2.2).

- Case where the observed factors M have daily frequency:

$$Y_{md} = A + B_M M_{md} + B_\theta \theta_{md} + \sigma u_{md},\ u_{md} \sim N(0, I_n), \tag{2.2}$$
$$M_{md} = \mu_M + \Phi_{MM} M_{m-1d} + \Phi_{M\theta}\theta_{m-1d} + \Sigma_{MM}\epsilon_{md}^M,\ \epsilon_{md}^M \sim N(0, I_p),$$
$$\theta_{md} = \mu_\theta + \Phi_{\theta M} M_{m-1d} + \Phi_{\theta\theta}\theta_{m-1d} + \Sigma_{\theta M}\epsilon_{md}^M + \Sigma_{\theta\theta}\epsilon_{md}^\theta,\ \epsilon_{md}^\theta \sim N(0, I_q).$$

- Case where observed factors M have monthly frequency:

$$Y_{md} = A + B_M M_m + B_\theta \theta_{md} + \sigma u_{md},\ u_{md} \sim N(0, I_n), \tag{2.3}$$
$$M_m = \mu_M + \Phi_{MM} M_{m-1} + \Phi_{M\theta}\theta_{m-1d} + \Sigma_{MM}\epsilon_{md}^M,\ \epsilon_{md}^M \sim N(0, I_p),$$
$$\theta_{md} = \mu_\theta + \Phi_{\theta M} M_{m-1} + \Phi_{\theta\theta}\theta_{m-1d} + \Sigma_{\theta M}\epsilon_{md}^M + \Sigma_{\theta\theta}\epsilon_{md}^\theta,\ \epsilon_{md}^\theta \sim N(0, I_q),$$

A monthly model is estimated considering only the set of information relative to a selected day d, or, alternatively, monthly averages. We, instead, take into account all days in a joint way, under the assumption that there are d independent models, one for each day d, and that all models share the same parameters. In the case the observed factors M have monthly frequency, the estimator must be corrected for the daily replications of the monthly factor. It turns out that only one step of the MCMC algorithm is altered, as discussed in the Inference section.

The models estimated here are the standard and the unrestricted versions of the affine (na) and dynamic Nelson-Siegel (ns) models (see chapter 4).

2.1. **Identification.** The econometric identification of the affine and unrestricted affine models is established by setting $\Phi^{\star}_{\theta M} = 0$, $\Phi^{\star}_{\theta\theta}$ upper triangular, $\Phi_{\theta\theta}$ lower triangular and $\delta_\theta = 1$. For ns and lg case, the T_L operator is constrained by $\beta = I$ in order to preserve the factor weights. Using α, we require that the last q (the number of latent factors) lines of B_M be constituted of zeroes.

In the common factor (cf) case, the chosen conditions are: $\Sigma_{\theta\theta} = I$, $\Sigma_{\theta M} = 0$ and $\Phi_{\theta\theta}$ lower triangular.

3. Inference

All models are estimated via MCMC. The full information likelihood (see Johannes and Polson, 2003) is given by

$$f(Y|\theta, M, \Psi) = \prod_t f(Y_t|\theta_t, M_t, \Psi)$$
$$= \prod_t \exp\left\{-1/2\left[\ln|\sigma\sigma'| + (Y_t - A - BX_t)'(\sigma\sigma')^{-1}(Y_t - A - BX_t)\right]\right\}.$$

The marginal likelihood is given by

$$f(Y|M, \Psi) = \prod_t f(Y_t|D_{t-1}) = \prod_t \exp\left\{-1/2\left[|Q_t| + (Y_t - f_t)' Q_t^{-1}(Y_t - f_t)\right]\right\},$$

where D_t represents all observed information up to t and Q_t, f_t are defined by Kalman filter equations seen in chapter 1.

For all the specifications, the estimation involved 7500 iterations of the algorithm are evaluated, of which the first 4500 are discarded (burn-in). The empirical distribution of the parameters is calculated using 1 every 3 of the last 3000 iterations to decrease the serial correlation of the Markov chain. Also, we assess the convergence of the chains through the Gelman-Rubin diagnostics.

In the appendix we present tests using simulated data. We show that the Monte Carlo chains converge under the Gelman-Rubin criterion, and that the true parameters remained within the confidence intervals of the parameter distributions generated by the chains.

We now explain our model that mixtures daily and monthly data. In the usual solution, the estimator uses "aggregated" or monthly data. For the case with daily yield and monthly macro, we have:

$$\begin{aligned} Y_m &:= \textstyle\sum_d Y_{md}/D, \\ Y_m &= A(.) + B(.)X_m + \sigma u_m, \\ X_m &= \mu + \Phi X_{m-1} + \Sigma \epsilon_m \end{aligned}$$

That is, monthly averages of the yield series are used for the estimation. An alternative is to use the yield for a selected day $\bar{d}$ of the month for every month, $Y_m = Y_{m\bar{d}}$.

Next, we transform a daily model into a sum of monthly models for each day d, whose likelihood is $L(\theta_d|D_d)$, where D_d is the information set associated to day d. We assume independence among the residues of each model, which allow the partition of the full likelihood into day-d likelihoods,

$$L(\theta|D) = \sum_{d=1}^{21} L(\theta|D_d).$$

The sum has 21 terms, the average number of business days in a month. In this way, we turn a daily model into 21 monthly models that are jointly estimated. The estimations are made under the assumption that

$$\theta^d = \theta.$$

This procedure can be used even for daily observed macro factors. In this case, the pooled version will be more precise than monthly models, because of the decreased variance of the estimator.

In the case the observed state factor has monthly frequency, we replicate the observed factor 21 times for each day of the month. The new estimator has the same mean as the aggregated estimator, except that the variance must be corrected to account for the introduced repetitions. However, only sub-problem 1 of the MCMC algorithm (see chapter 4) is affected by the replicated macro data, specifically the parameter Φ of the transition equation of the observed variables. We correct the variance of the distribution of sub-problem 1 as follows.

Estimator for day d:

$$E(\hat{\Phi}_{md}) = \begin{pmatrix} M_m'M_m & M_m'\theta_{md} \\ (M_m'\theta_{md})' & \theta_{md}'\theta_{md} \end{pmatrix}^{-1} \begin{pmatrix} M_m'M_{m+1} \\ \theta_{md}'M_{m+1} \end{pmatrix}.$$

Pooled estimator: replicate $M_{md} = M_m \forall d$. Stacking the equations for the $d = 1..D$ days of the month, we have:

$$\begin{aligned} E(\hat{\Phi}_m) &= \begin{pmatrix} M_{md}'M_{md} & M_{md}'\theta_{md} \\ (M_{md}'\theta_{md}) & V_d \end{pmatrix}^{-1} \begin{pmatrix} M_{md}'M_{m+1d} \\ \theta_{md}'M_{m+1d} \end{pmatrix} \\ &= \begin{pmatrix} DM_m'M_m & DM_m'\overline{\theta}_m \\ (DM_m'\overline{\theta}_m)' & D\overline{V} \end{pmatrix}^{-1} \begin{pmatrix} M_{md}'M_{m+1d} \\ \theta_{md}'M_{m+1d} \end{pmatrix} \\ &= \begin{pmatrix} M_m'M_m & M_m'\overline{\theta}_m \\ (M_m'\overline{\theta}_m)' & \overline{V} \end{pmatrix}^{-1} \begin{pmatrix} M_m'M_{m+1} \\ \overline{\theta}_m'M_{m+1} \end{pmatrix} = E(\hat{\Phi}_m) \end{aligned}$$

where $\hat{\Phi}_m$ is the monthly estimator and

$$V_d = \theta_{md}'\theta_{md}, \overline{V} = \sum_d V_d/D, \overline{\theta}_m = \sum_d \theta_{md}/D.$$

That is, the mean of the pooled and of the aggregated estimators are equal. However, the variances must be corrected by the factor D:

$$\begin{aligned} V(\hat{\Phi}_m) &= \sigma^2 \begin{pmatrix} DM_m'M_m & DM_m'\overline{\theta}_m \\ (DM_m'\overline{\theta}_m)' & D\overline{V} \end{pmatrix}^{-1} \\ &= D^{-1}\sigma^2 \begin{pmatrix} M_m'M_m & M_m'\overline{\theta}_m \\ (M_m'\overline{\theta}_m)' & \overline{V} \end{pmatrix}^{-1} = D^{-1}V(\hat{\Phi}_m), \end{aligned}$$

where

$$\sigma^2 = \sum_{md}(M_{m+1} - \Phi_{MM}M_m - \Phi_{M\theta}\theta_{md})^2/\#(md).$$

3.1. Information Criterion. The in-sample criteria used to order the models are the posterior predictive loss (ppl) and the deviance of information criterion (dic), discussed in chapter 4.

4. Empirical Analysis

To situate our work in the literature, we brie‡y discuss the results of 3 articles having similar objectives. The article by Diebold et al. (2006) is a monthly model with 3 latent factors, estimated with data from 01/1987 to 12/2001 consisting of the U.S. yield curve (YC), output gap and in‡ation. Using a dynamic Nelson-Siegel model with B_M = 0 and where shocks are identi…ed such that latent factors are more exogenous macro indicators, they conclude that: 1) the macro shocks explain only 1% of the level and slope movements 12 months after the shock; 2) the e¤ect of the output gap shock on the level and slope is positive, and the e¤ect of the in‡ation shock is null.

On a more macroeconomic view, Rudebusch and Wu (2003) discuss "structural" restrictions on the relation between macro indicators (in‡ation and output gap) and latent factors. Those restrictions allow the interpretation of the latent factors as monetary policy factors. It is a monthly a¢ne model using data from 01/1988 to 12/2000. Their results are: 1) the economic indicators do not explain much of the YC; 2) the in‡ation and output gap shocks increase the slope but explain little of the level movements.

Finally, Ang et al. (2007) use an a¢ne model to estimate U.S. quarterly data from 06/1952 to 12/2004. They allow the transition of the macro variables – output gap, in‡ation – to follow an autoregressive process with order greater than 1. Implementation of di¤erent monetary policy rules, including backward and forward-looking Taylor rules, are discussed. Impulse response functions of the YC are not presented, but it is documented that the output gap and in‡ation respond to about 29% and 38% of the level of the rates, and 73% and 20% of the slope. Therefore, their article shows signi…cant macro e¤ects on the YC.

We now show the impulse response functions and variance decompositions of the models. We compute the response of the yield curve and macro factors to state variable shocks. In the notation of chapter 3, the responses are denoted IRFY and IRFM. As proved there, these responses are independent of the latent factor identi…cation, but are dependent on the choice of the ordering of the factors. Here, we impose that macro factors are more exogenous than latent factors.

In the last chapter, we have seen that by the information criterion, the unrestricted Nelson-Siegel model is the best choice in most cases. For this reason, in all this section we have used this model. We use 3 latent factors for U.S. and 2 for Brazil, since in the last case the longest maturity is 3 years.

The data used here are the same as in chapter 4, and the models are estimated using one of 2 orderings of exogeneity of the state factors. Order 1: output gap, in‡ation, level, slope. Order 2: output gap, in‡ation, slope, level.

Table 1 reports the variance decompositions of the Brazilian case using the di¤erent unrestricted Nelson-Siegel versions. The macro factors explain relatively less of the yields movements than vice-versa. Also, as expected, the standard deviations of the pooled version are smaller.

Table 1. Brazil. Variance decompositions 6 months after shock. Upper half: on each column, contributions of shocks (s-gap, s-in‡, s-lev and s-slop) on the variation of the variables. Lower half: standard deviations of upper half.

	resp: pooled order 1				resp: aggreg. order 1				resp: aggreg. order 2			
shock	gap	in‡	shrt	long	gap	in‡	shrt	long	gap	in‡	shrt	long
s-gap	0.69	0.05	0.03	0.01	0.59	0.07	0.12	0.03	0.70	0.05	0.03	0.01
s-in‡	0.07	0.70	0.08	0.05	0.08	0.64	0.11	0.12	0.06	0.70	0.08	0.05
s-lev	0.05	0.22	0.32	0.93	0.09	0.25	0.30	0.83	0.22	0.02	0.78	0.12
s-slop	0.19	0.03	0.57	0.01	0.24	0.04	0.47	0.03	0.02	0.23	0.11	0.82
standard deviation												
s-gap	0.07	0.03	0.03	0.01	0.13	0.04	0.12	0.03	0.07	0.04	0.02	0.01
s-in‡	0.04	0.08	0.04	0.02	0.05	0.08	0.11	0.07	0.04	0.08	0.04	0.01
s-lev	0.04	0.08	0.07	0.02	0.08	0.08	0.18	0.09	0.07	0.02	0.06	0.01
s-slop	0.06	0.02	0.06	0.01	0.10	0.03	0.16	0.03	0.02	0.08	0.05	0.02

Next, we present the impulse response functions of the unrestricted model in …gure I. The aggregated and pooled versions are compared in each graph. The rows present the responses of the in‡ation, short rate and slope to industrial output gap, in‡ation, level and slope shocks.

Instead of plotting the mean response, we plot the one deviation above response and the one deviation below. In some cases the impulse response con…dence interval of the pooled version is not contained on the con…dence interval of the aggregated model, indicating that the pooled version may contain information that cannot be discarded.

Table 2 shows that macro factors constitute important factors to explain the yield curve, specially the short yield. Likewise, the monetary factors explain an important part of the macro factors, specially the industrial output gap.

Table 2. U.S. Variance decompositions 12 months after shock. Upper half: on each column, contributions of shocks (s-gap, s-in‡, s-lev, s-slop, s-curv) on the variation of the variables. Lower half: standard deviations of upper half.

	resp: pooled order 1				resp: aggreg. order 1				resp: aggreg. order 2			
shock	gap	in‡	shrt	long	gap	in‡	shrt	long	gap	in‡	shrt	long
s-gap	0.68	0.00	0.17	0.04	0.61	0.00	0.16	0.05	0.61	0.00	0.16	0.05
s-in‡	0.01	0.96	0.04	0.05	0.01	0.94	0.06	0.07	0.01	0.94	0.06	0.07
s-lev	0.18	0.03	0.38	0.87	0.20	0.04	0.38	0.81	0.01	0.01	0.17	0.21
s-slop	0.09	0.01	0.24	0.01	0.10	0.01	0.19	0.01	0.17	0.00	0.06	0.15
s-curv	0.05	0.01	0.17	0.03	0.07	0.01	0.21	0.06	0.20	0.03	0.56	0.52
standard deviation												
s-gap	0.08	0.00	0.02	0.01	0.08	0.00	0.06	0.02	0.08	0.00	0.05	0.02
s-in‡	0.01	0.02	0.01	0.01	0.02	0.03	0.04	0.04	0.02	0.03	0.04	0.03
s-lev	0.06	0.02	0.03	0.01	0.06	0.02	0.07	0.05	0.01	0.01	0.04	0.04
s-slop	0.03	0.01	0.02	0.00	0.03	0.01	0.06	0.01	0.04	0.01	0.02	0.03
s-curv	0.03	0.01	0.03	0.01	0.04	0.01	0.08	0.04	0.09	0.03	0.10	0.07

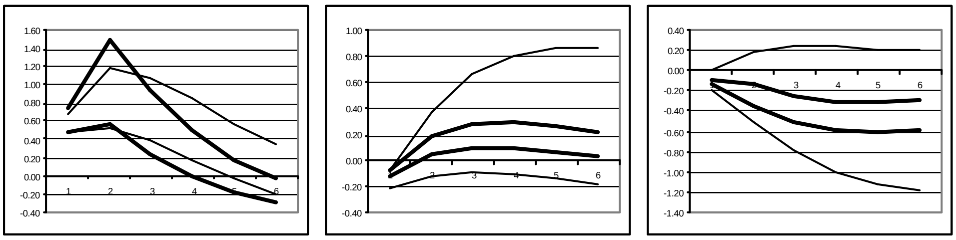
The pooled and the aggregated versions results are very similar. This result indicates that the U.S. series is long enough, and so monthly data already produces accurate results.

Continuing, we show in figure II the impulse response functions of the unrestricted Nelson-Siegel model with U.S. data. In all cases, the one deviation above and below impulse response functions of the pooled version is contained in the aggregated model confidence interval. This means that both versions contain equivalent information, the only difference being that the pooled is more accurate in some cases.

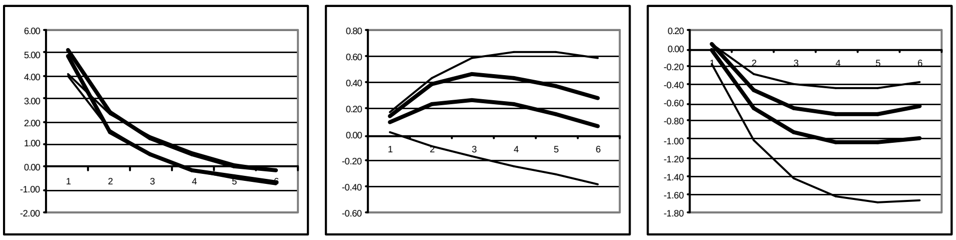
Figure I. Brazil. Impulse response. Thick line: pooled version. Thin line: aggregated version. We plot 95% confidence bands around responses.

1^{st} column: inflation resp. 2^{nd} col.: short rate resp. 3^{rd} col.: slope resp.

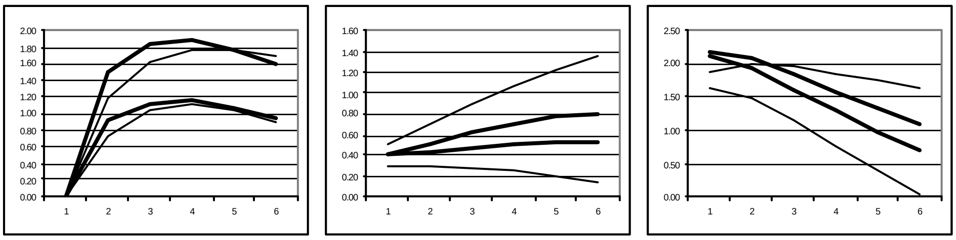
output gap shock



inflation shock



level shock



slope shock

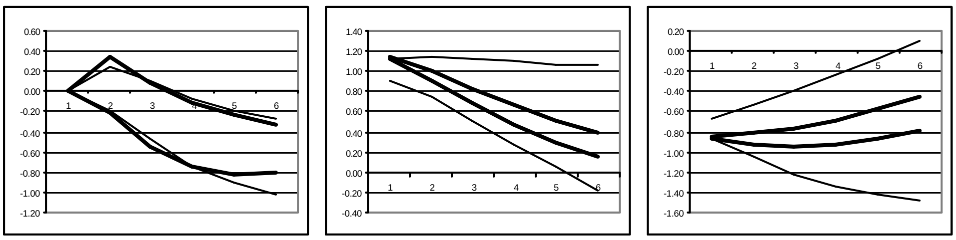
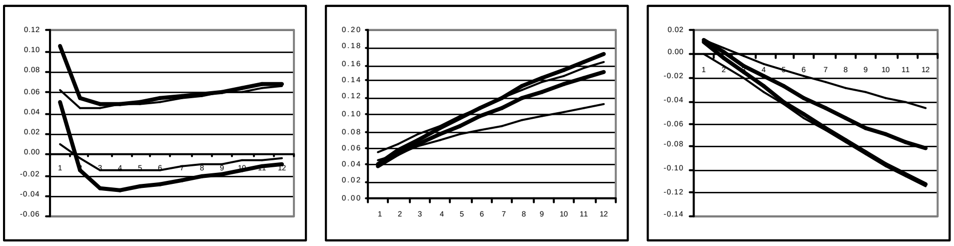


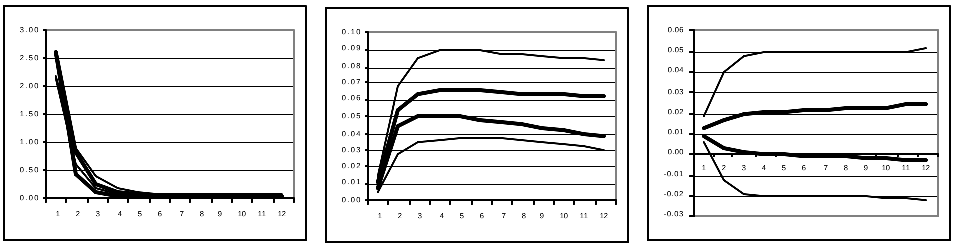
Figure 11. U.S. Impulse response. Thick line: pooled version. Thin line: aggregated version. We plot 95% con…dence bands around responses.

1^{st} column: in‡ation resp. 2^{nd} col.: short rate resp. 3^{rd} col.: slope resp.

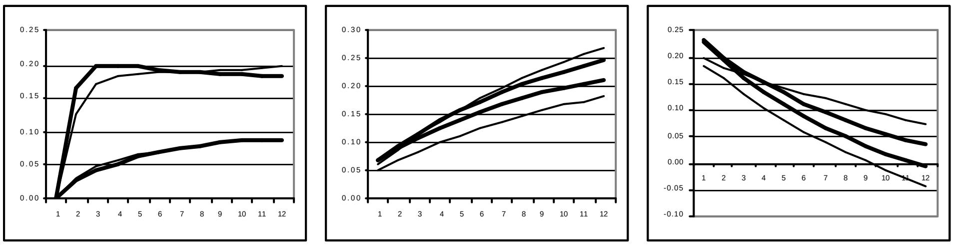
gap shock



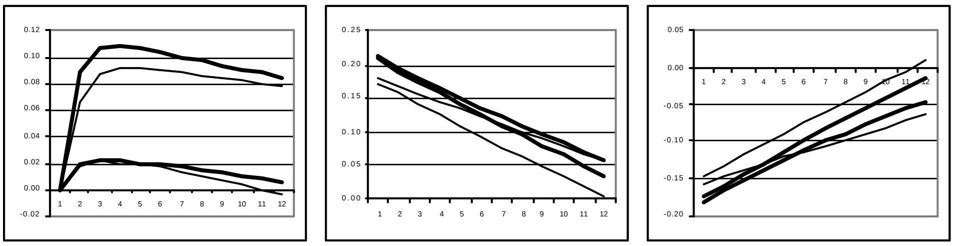
in‡ation shock



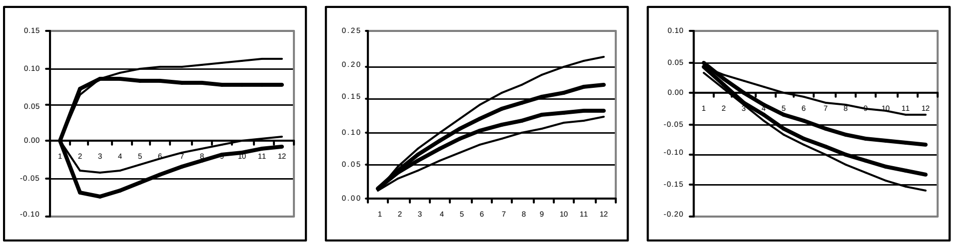
level shock



slope shock



curvature shock



5. Change of regime: Taylor rule switching model and sub-sample

In order to allow the possibility of changes of regime associated to economic events, we propose two extensions of the main models. We are speci…cally interested in the change of the chairman or president of the Monetary Authority.

In approximate terms, we can consider that the Monetary Authority reaction function is the short-run impulse response of the short rate, which is given by

$$\frac{\partial y_t^1}{\partial \epsilon_t} = \frac{\partial y_t^1}{\partial \theta_t}\frac{\partial \theta_t}{\partial \epsilon_t} = \delta^{\text{I}}\,\S. \tag{5.1}$$

This impulse response is calculated for the a¢ne model in chapter 1. However, it is proved there that we may choose the identi…cation scheme that …xes $\delta = (\delta^M, \delta^\theta)^{\text{I}} = (0,1)^{\text{I}}$. In this case, monetary policy changes translates into § changes.

Motivated by this, we propose the following "Taylor rule" switching model:

$$Y_t = A + B_\theta \theta_t + \sigma u_t, u_t \text{ » } N(0, I_n),$$

$$X_t = \copyright X_{t\,¡\,h} + \S_t \epsilon_t,\ \ \epsilon_t \text{ » } N(0, I_p),$$

$$\S_t = \S_0 1_{[t<t_0]} + \S_1 1_{[t\,¸\,t_0]}.$$

This extension is estimated with the same data as in previous sections.

A simpler way to analyze change would be to consider and compare estimations using sub-samples corresponding to the possible changes of regime, but this requires enough data.

For Brazil, the event that marked the change point was the presidential change from president Cardoso (FHC) to president Lula. The sample contains about 4 years of each period. In the case of U.S., we considered more sensible to mark the event of the change of the Federal Reserve's chair. Our sample contain the 18 years period of Greenspan, and 2 years of Greenspan.

The variance decompositions below compare the responses before and after those changes.

Table 3. FHCnLula. Pooled. Variance decompositions 6 months after shock. Upper half: contributions of shocks (s-gap, s-in‡, s-lev, s-slop, s-curv). Lower half: standard deviations.

	FHC				Lula			
	gap	in‡	shrt	long	gap	in‡	shrt	long
s-gap	0.68	0.05	0.02	0.01	0.86	0.07	0.17	0.05
s-in‡	0.07	0.68	0.12	0.03	0.07	0.79	0.03	0.17
s-slop	0.05	0.25	0.33	0.95	0.01	0.13	0.24	0.76
s-lev	0.20	0.02	0.52	0.01	0.07	0.01	0.56	0.01
standard deviation								
s-gap	0.06	0.02	0.01	0.00	0.04	0.04	0.05	0.02
s-in‡	0.03	0.06	0.03	0.01	0.04	0.04	0.02	0.03
s-slop	0.04	0.06	0.04	0.01	0.01	0.03	0.03	0.04
s-lev	0.05	0.01	0.03	0.00	0.02	0.01	0.04	0.00

Table 4. GreenspannBernanke. Pooled. Variance decompositions 12 months after shock. Upper half: contributions of shocks (s-gap, s-in‡, s-lev, s-slop, s-curv). Lower half: standard deviations.

	Greenspan				Bernanke			
	gap	in‡	shrt	long	gap	in‡	shrt	long
s-gap	0.77	0.02	0.16	0.03	0.63	0.05	0.02	0.02
s-in‡	0.01	0.95	0.04	0.07	0.05	0.93	0.04	0.16
s-slop	0.11	0.01	0.28	0.85	0.26	0.01	0.49	0.77
s-lev	0.06	0.00	0.27	0.01	0.05	0.00	0.43	0.04
s-curv	0.06	0.01	0.26	0.04	0.01	0.00	0.03	0.01
standard deviation								
s-gap	0.06	0.01	0.02	0.01	0.09	0.03	0.01	0.01
s-in‡	0.01	0.03	0.01	0.01	0.03	0.04	0.01	0.03
s-slop	0.04	0.02	0.03	0.01	0.10	0.02	0.03	0.03
s-lev	0.02	0.01	0.02	0.00	0.03	0.01	0.03	0.01
s-curv	0.03	0.02	0.03	0.01	0.01	0.00	0.01	0.00

6. Conclusion

The relation between the YC and the macro variables have been examined by diverse authors that propose di¤erent approaches to relate them, analyzing in general developed markets, estimating with monthly or quarterly data.

Here we make a comprehensive study of this relation, examining in parallel a mature market (U.S.) and a emerging market (Brazil), assessing di¤erent approaches for the speci…cation of the latent factors, using a Bayesian inference method (MCMC), and proposing methodological improvements.

The …nance models usually use daily or weekly samples, and macro models usually use low frequency data. The most straightforward solution to this discrepancy is to discard (or take averages of) the …nancial data to adequate to the macro variables frequency.

This solution can be troublesome for emerging markets such as the Brazilian, which in general dispose of shorter samples of stable monetary and economic regimes. To deal with it, we propose a modi…cation of the monthly model that uses all information available from the …nancial data.

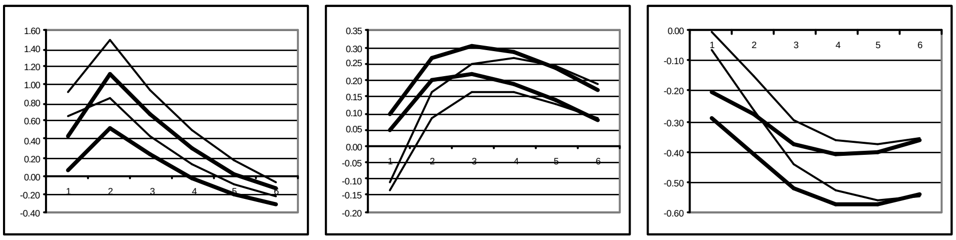
We showed that for the Brazilian data, the "pooled" model produced more accurate impulse response functions.

Finally, we considered a changes of regime model.

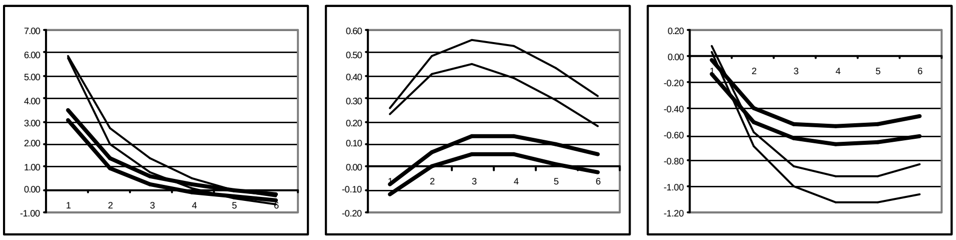
Figure III. FHCnLula. Pooled. Impulse response. Thick line: Lula. Thin line: FHC. We plot 95% con…dence bands around responses.

1^{st} column: in‡ation resp. 2^{nd} col.: short rate resp. 3^{rd} col.: slope resp.

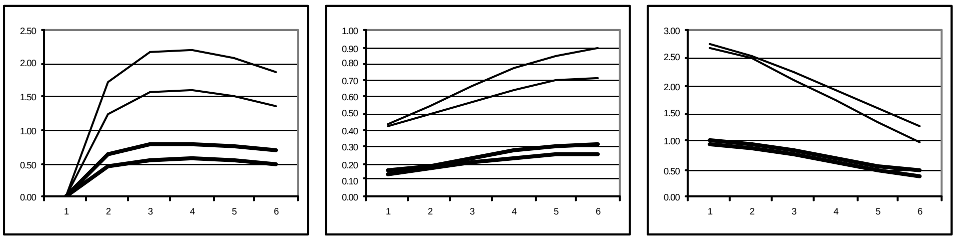
gap shock



in‡ation shock



level shock



slope shock

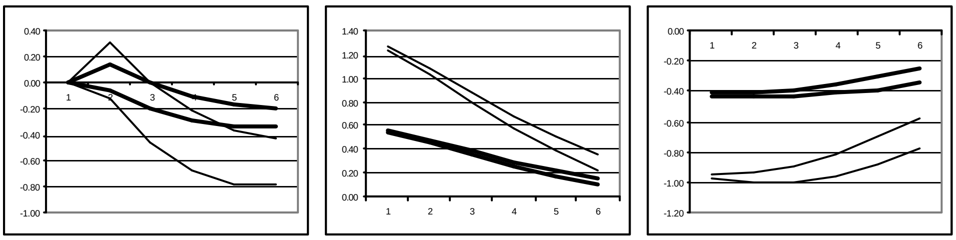
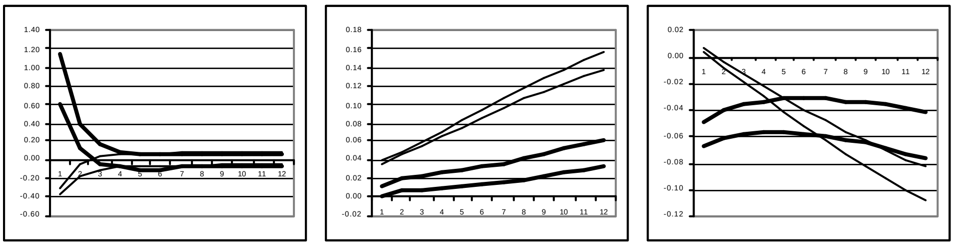


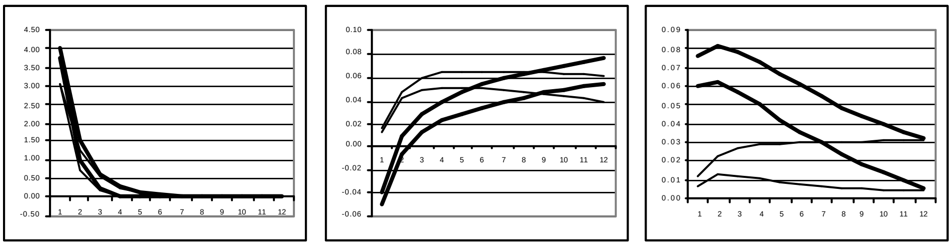
Figure IV. GreenspannBernanke. Pooled. Impulse response. Thick line: Bernanke. Thin line: Greenspan. We plot 95% con…dence bands around responses.

1^{st} column: in‡ation resp. 2^{nd} col.: short rate resp. 3^{rd} col.: slope resp.

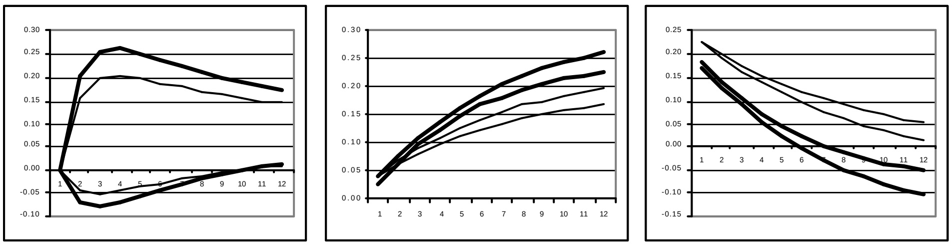
gap shock



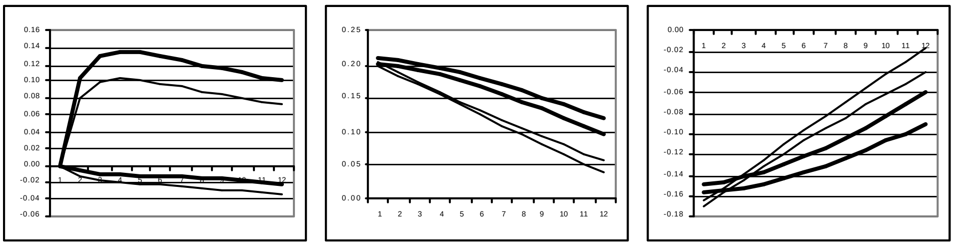
in‡ation shock



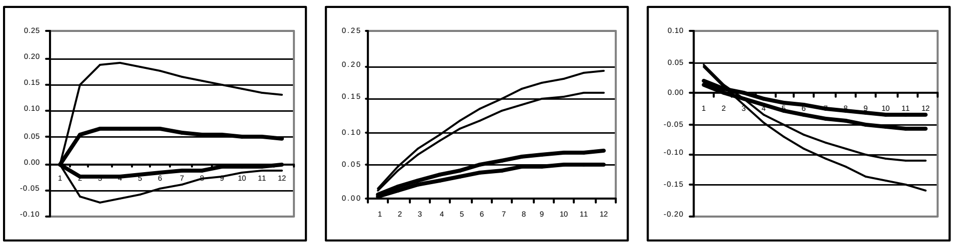
level shock



slope shock



curvature shock



APPENDIX A

Convergence

1. Description of Gelman-Rubin test

Below, we give the formulas associated to the Gelman-Rubin statistics:

Within chain variance:

$$W = \frac{1}{m(n-1)} \sum_{j=1} \sum_{i=1} (x_j^i - \overline{x}_j)^2.$$

Among chains variance:

$$B = \frac{n}{m-1} \sum_{j=1} (x_j - \overline{x})^2.$$

Estimated variance:

$$\hat{V}(x) = (1 - \frac{1}{n})W + \frac{1}{n}B.$$

Gelman-Rubin statistics:

$$R = \frac{\hat{V}(\theta)}{W}.$$

The Gelman-Rubin statistics R are applied to multiple chains. It indicates convergence when R is close to 1, below a certain critical level, for example, 1.1.

Details can be found in Robert and Casella (2005).

2. Simulation exercise

To validate our MCMC estimation algorithm, we simulate series using the unrestricted Nelson-Siegel and affine models, and test the performance of the estimator on the series.

We show in the following that the algorithm successfully recovers the true parameters. We simulate a sample of 1000 observations of a term structure with the same maturities as the U.S. market. We show that the MCMC estimates are close to the true parameters. Moreover, the true parameters stay within the confidence intervals of the estimated data.

Moreover, we apply the convergence diagnostics for the parameter chains generated to estimate the above parameters.

For the Nelson-Siegel model, the results are given in Table A1.

We further plot the Markov chains produced by the MCMC algorithm for various parameters. For each parameter, 3 independent chains were run 3000 times. The results are given by Figures A1 to A4.

In all cases the chains converged to its invariant measure around the true parameter. We also compute the Gelman-Rubin statistics for the chains (some are omitted).

In Table A2, we plot the Markov chains produced by the MCMC algorithm for the unrestricted a¢ne case. For each parameter, 3 independent chains were run 3000 times. In all cases the chains converged to its invariant measure around the true parameter. Gelman-Rubin statistics are given (some are omitted), as well as the convergence diagnostics of the parameter chains.

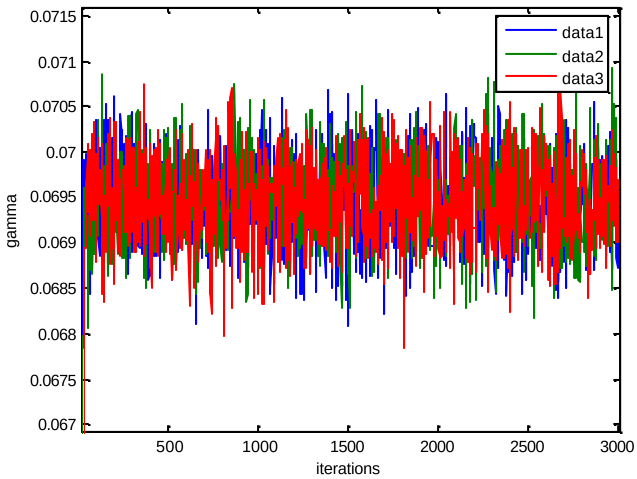
Table A1. Nelson-Siegel with 3 factors. True parameters, estimated parameters (mean and standard deviation of the parameter chains).

	True			Mean			St. Dev.		
©	0.9	0.0	-0.1	0.87	0.02	-0.07	0.02	0.02	0.02
	0.1	0.88	-0.1	0.12	0.87	-0.1	0.01	0.01	0.01
	0.0	0.0	0.98	-0.01	0.00	0.98	0.01	0.01	0.01
V ¤ 10^5	3.6	-0.6	0.6	4.00	-0.34	0.46	0.09	0.03	0.03
	-0.6	1.0	-0.4	-0.34	1.01	-0.41	0.03	0.02	0.01
	0.6	-0.4	1.8	0.46	-0.41	1.57	0.03	0.01	0.02

	True	Mean	St. Dev.
γ ¤ 10^2	7	6.99	0.05

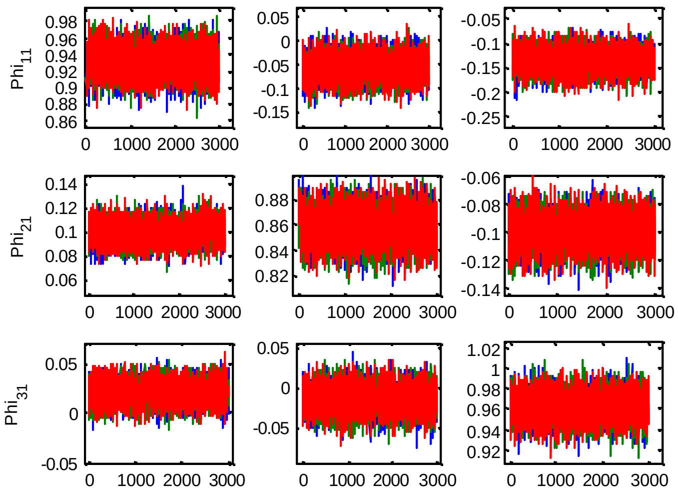
A ¤ 10^2	True	5	5	5	5	5	5	5	5
	Estim	5.0	5.0	4.9	4.9	5.0	5.0	5.0	5.0
	St. Dev.	-	-	-	-	-	-	-	-
σ_u ¤ 10^4	True	4	4	4	4	4	4	4	4
	Estim	3.9	4.2	4.1	4.0	4.0	4.0	4.3	3.9
	St. Dev.	0.2	0.1	0.1	0.1	0.2	0.2	0.1	0.2

Figure A1. Nelson-Siegel with 3 factors.
Chains of parameter γ.



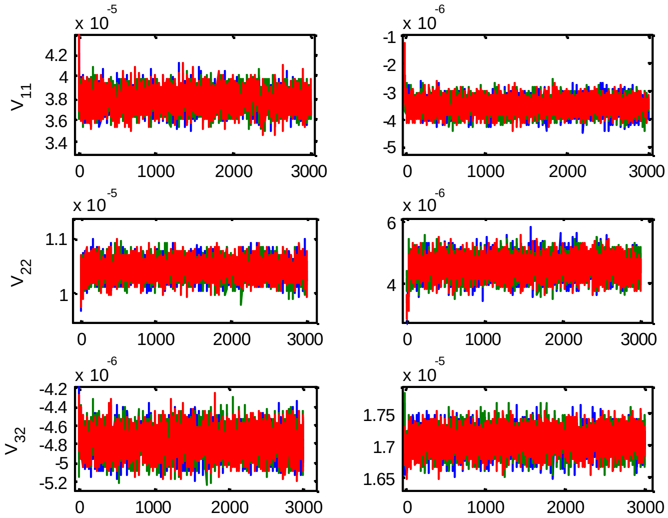
Gelman-Rubin test:
$R(\gamma)$ = [1.00, 1.00, 1.00]

Figure A2. Nelson-Siegel with 3 factors.
Chains of parameter matrix ©.



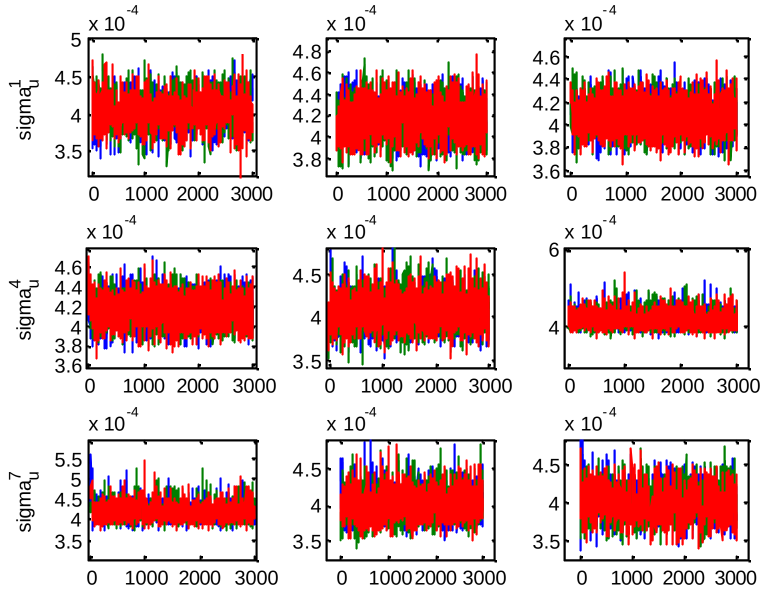
Gelman-Rubin test:
$R(©_{11})$ = [1.00, 1.00, 1.00]
$R(©_{22})$ = [1.00, 1.00, 1.00]
$R(©_{33})$ = [1.00, 1.00, 1.00]

Figure A3. Nelson-Siegel with 3 factors.
Chains of parameter matrix V.



Gelman-Rubin test:
$R(V_{11})$ = [1.00, 1.00, 1.00]
$R(V_{22})$ = [1.00, 1.00, 1.00]
$R(V_{33})$ = [1.00, 1.00, 1.00]

Figure A4. Nelson-Siegel with 3 factors.
Chains of parameter vector σ.



Gelman-Rubin test:
$R(\sigma_u^1)$ = [1.00, 1.00, 1.00]
$R(\sigma_u^4)$ = [1.00, 1.00, 1.00]
$R(\sigma_u^7)$ = [1.00, 1.00, 1.00]

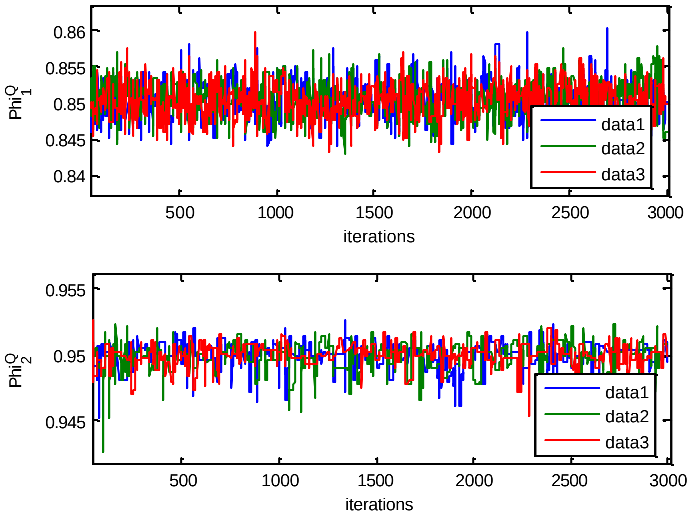
Table A2. Unrestricted A¢ne with 2 factors. True parameters, estimated parameters (mean and standard deviation of the parameter chains).

param	true value		est. mean		est. std. dev.	
©	0.90	-0.05	0.90	-0.03	0.02	0.02
	0.10	0.98	0.10	0.97	0.02	0.02
©^{Q}	0.85	0.00	0.849	0.000	0.003	0.000
	0.00	0.95	0.000	0.950	0.000	0.002
V ¤ 10^5	1.0	-0.2	1.05	-0.17	0.01	0.01
	-0.2	0.4	-0.17	0.39	0.01	0.01

A ¤ 100	true value	4.82	4.57	4.73	4.89	5.24	5.43	5.73	5.97	6.10
	est. mean	4.71	4.46	4.62	4.79	5.16	5.37	5.69	5.94	6.08
	est. st. dev.	-	-	-	-	-	-	-	-	-

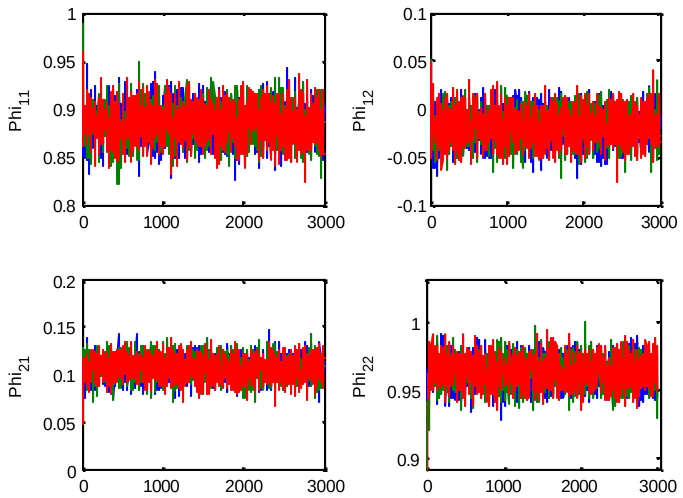
σ_u ¤ 10^4	true value	4	4	4	4	4	4	4	4	4
	est. mean	3.7	4.5	4.8	5.8	6.7	6.5	5.8	5.1	4.6
	est. st. dev.	0.3	0.3	0.9	1.8	2.2	2.2	1.8	1.4	1.0

Figure A5. Unrestricted a¢ne with 2 factors.
Chains of parameter matrix $©^{Q}$.



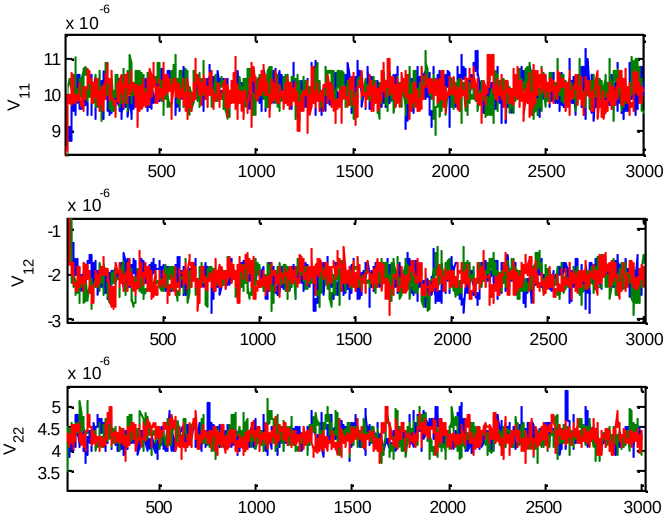
Gelman-Rubin statistics:
$R(©^{Q}_{1})$ = [1.00, 1.02, 1.03]
$R(©^{Q}_{2})$ = [1.00, 1.05, 1.03]

Figure A6. Unrestricted a¢ne with 2 factors.
Chains of parameter matrix ©.



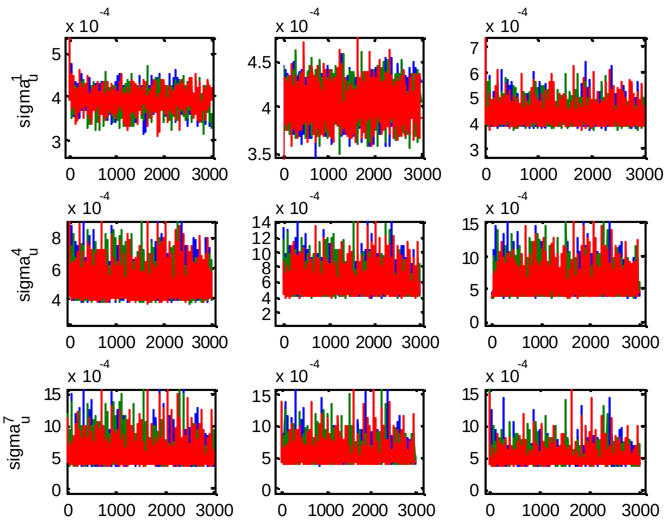
Gelman-Rubin statistics:
$R(©_{11})$ = [1.00, 1.01, 1.00]
$R(©_{22})$ = [1.00, 1.01, 1.01]

Figure A7. Unrestricted a¢ne with 2 factors.
Chains of parameter matrix V.



Gelman-Rubin statistics:
$R(V_{11})$ = [1.00, 1.02, 1.01]
$R(V_{22})$ = [1.00, 1.02, 1.01]

Figure A8. Unrestricted a¢ne with 2 factors.
Chains of parameter vector σ^u.



$R(\sigma_1^u)$ = [1.00, 1.00, 1.03]
$R(\sigma_4^u)$ = [1.00, 1.00, 1.00]
$R(\sigma_7^u)$ = [1.00, 1.00, 1.00]

APPENDIX B

References

(1) Almeida, C., Duarte, A., Fernandes, C. (1998) Decomposing and simulating the movements of term structures in emerging eurobonds markets, Journal of Fixed Income, 1, 21–31.

(2) Almeida, C., Gomes, R., Leite, A., Vicente, J. (2007A) Movimentos da estrutura a termo e critérios de minimização do erro de previsão em um modelo paramétrico exponencial, Working paper series 146, Central Bank of Brazil.

(3) Almeida, C., Gomes, R., Leite, A., Vicente, J. (2007B) Does curvature enhance forecasting? Working paper series 155, Central Bank of Brazil.

(4) Almeida, C., Vicente, J. (2008) The role of no-arbitrage on forecasting: lessons from a parametric term structure model, forthcoming in Journal of Banking and Finance.

(5) Almeida, C. and J. Vicente (2009). Identifying volatility risk premia from ...xed income Asian options. Journal of Banking and Finance, 33, p. 652-661.

(6) Amato, J., Luisi, M. (2006) Macro factors in the term structure of credit spreads, BIS Working Papers No. 203.

(7) Ang, A., Dong, S., Piazzesi, M. (2007) No-arbitrage Taylor rules, NBER Working Paper 13448.

(8) Ang, A., Piazzesi, M. (2003) A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, Journal of Monetary Economics, 50, 745–787.

(9) Banerjee, S., Carlin, B., Gelfand, A. (2004) Hierarchical Modeling and Analysis. for Spatial Data. CRC Press/Chapman Hall.

(10) Berndt, A., R. Douglas, D. Du¢e, M. Ferguson, and D. Schranzk (2004). Measuring default risk premia from default swap rates and EDFs. Working Paper, Stanford University.

(11) Black, F. and Cox, J. (1976) Valuing corporate securities: some e¤ects of bond indenture provisions, Journal of Finance, 31(2), 351–367.

(12) Black, F. and Scholes, M. (1973) The pricing of options and corporate liabilities, Journal of Political Economy, 81(3), 637–654.

(13) Bogdanski, J., Tombini, A., Werlang, S. (2000) Implementing In‡ation Target in Brazil, Working Paper Series 1, Central Bank of Brazil.

(14) Chen, L.-H. (2001) In‡ation, real short-term interest rates, and the term structure of interest rates: a regime-switching approach, Applied Economics, 33(3), 393-400.

(15) Chen, R., Scott, L. (1993) Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates, Journal of Fixed Income, 3, 14–31.

(16) Collin-Dufresne, P., Goldstein, R., Jones, C. (2006) Identi…cation of maximal a¢ ne term structure models. Working paper.
(17) Cox, J., Ingersoll, J., Ross, S. (1985) A theory of the term structure of interest rates. Econometrica, 53: 385-407.
(18) Dai, Q., Philippon, T. (2004) Government de…cits and interest rates: a no-arbitrage structural VAR approach, Working paper.
(19) Dai, Q., Singleton, K. (2000) Speci…cation analysis of term structure of interest rates, Journal of Finance, 55, 1943–78.
(20) Davidson, R., MacKinnon, J. (1993) Estimation and inference in econometrics. New York: Oxford University Press.
(21) Dewachter, H., Lyrio, M. (2006) Macro Factors and the Term Structure of Interest Rates. Journal of Money, Credit and Banking, 38(1), 119–140.
(22) Diebold, F., Li, C. Forecasting the term structure of government bond yields. Journal of Econometrics, 130, 337-364.
(23) Diebold, F., Rudebusch, G., Aruoba, S. (2006) The macroeconomy and the yield curve: a dynamic latent factor approach, Journal of Econometrics, 131, 309–338.
(24) Diebold, F., Piazzesi, M., Rudebusch, G. (2005) Modeling bond yields in …nance and macroeconomics, American Economic Review, 95, N. 2, 415–20.
(25) Du¤ee, G. (2002). Term premia and interest rates forecasts in a¢ ne models. Journal of Finance, 57, p. 405-443.
(26) Du¢ e, D. (2001) Dynamic asset pricing theory, 3rd ed., Princeton: Princeton University Press.
(27) Du¢ e, D., Kan, R. (1996) A yield-factor model of interest rates, Mathematical Finance, 6, 379–406.
(28) Du¢ e, D., Pedersen, L., Singleton, K. (2003) Modeling sovereign yield spreads: a case study of Russian debt, Journal of Finance, 58, 119-159.
(29) Du¢ e, D. and Singleton, K. (1999) Modeling term structures of defaultable bonds, Review of Financial Studies, 12, 687–720.
(30) Fachada, P. (2001) In‡ation targeting in Brazil: reviewing two years of monetary policy 1999/00, Working paper series 25, Central Bank of Brazil.
(31) Fackler, P. (2000). Moments of a¢ ne di¤usions, Working paper, North Carolina State University.
(32) Fraga, A., Goldfajn, I., Minella, A. (2003) In‡ation targeting in emerging market economies, Working paper series 76, Central Bank of Brazil.
(33) Freitas, P., Muinhos, M. (2001) A simple model for in‡ation targeting in Brazil, Working paper series 18, Central Bank of Brazil.
(34) Gamerman, D., Lopes, H. (2006) Markov chain monte carlo: stochastic simulation for Bayesian inference, 2nd Ed. London: Chapman & Hall.
(35) Gelfand, A., Ghosh, S. (1998) Model choice: a minimum posteriori predictive loss approach. Biometrika, 85, 1-11.
(36) Ghezzi, P., Xu, D. (2002) From fundamentals to spreads. Deutsche Bank Global Markets Research.
(37) Giesecke, K. (2004) Credit risk modeling and valuation: an introduction, in Credit Risk Models and Management Vol. 2, (Ed.) D. Shimko, Riskbooks.

(38) Goldfajn, I., Werlang, S. (2000) The pass-through from depreciation to in‡ation: a panel study, Working paper series 5, Central Bank of Brazil.

(39) Goldman Sachs Global Economics Department (2007). BRICs and beyond.

(40) Hamilton, J. (1994) Time Series Analysis, Princeton University Press.

(41) Harvey, A. (1989) Forecasting, structural time series models and the Kalman …lter. Cambridge University Press: Cambridge.

(42) Hejazi, W. (2000) Yield spreads as predictors of industrial production: expectations on short rates or term premia?, Applied Economics, 32, 945-51.

(43) Hördahl, P., Tristani, O., Vestin, D. (2002) A joint econometric model of macroeconomic and term structure dynamics. European Central Bank Working Paper.

(44) Johannes, M., Polson, N. (2006) MCMC methods for continuous-time …nancial econometrics, in Handbook of Financial Econometrics (Ed.) Aït-Sahalia, Y, Hansen, L.

(45) Kalev, P., Inder, B. (2006) The information content of the term structure of interest rates, Applied Economics, 38(1), 33-45.

(46) Lando, D. (1998) On Cox processes and credit risky bonds, Review of Derivatives Research, 2(2), 99–120.

(47) Laurini, M., Hotta, L. (2008) Bayesian extensions to diebold-li term structure model, Ibmec working paper 74.

(48) Leippold, M., Wu, L. (2003) Estimation and design of quadratic term structure models, Review of Finance, 7(1), 47–73.

(49) Leland, H. (1994) Risky debt, bond covenants and optimal capital structure, Journal of Finance, 49 (4), 1213–1252.

(50) Leland, H., Toft, K. B. (1996) Optimal capital structure, endogenous bankruptcy and the term structure of credit spreads, Journal of Finance, 50, 789–819.

(51) Litterman, R., Scheinkman, J. (1991) Common factors a¤ecting bond returns, Journal Fixed Income, 1, 51–61.

(52) Longsta¤, F. (1989) A nonlinear general equilibrium model of the term structure of interest rates, Journal of Financial Economics 23, 195–224.

(53) Ludvigson, S., Ng, S. (2007) Macro factors in bond risk premia, The Review of Financial Studies, forthcoming.

(54) Kalev, P., Inder, B. (2006) The information content of the term structure of interest rates, Applied Economics, 38(1), 33-45.

(55) Merton, R. (1974) On the pricing of corporate debt: the risk structure of interest rates, Journal of Finance 29(2), 449–470.

(56) Minella, A., Freitas, P., Goldfajn, I., Muinhos, M. (2003) In‡ation targeting in Brazil: constructing credibility under exchange rate volatility, Working paper series 77, Central Bank of Brazil.

(57) Moench, E. (2008) Forecasting the yield curve in a data-rich environment: a no-arbitrage factor-augmented VAR approach, Journal of Econometrics, 146(1).

(58) Moody's (2008). Sovereign default and recovery rates, 1983-2007, Working Paper, Moody's Global Credit Research. Available at http://www.moodys.com/.

(59) Muinhos, M., Alves, S. (2003) Medium-size macroeconomic model for the Brazilian economy, Working paper series 64, Central Bank of Brazil.

(60) Moreira, A., Rocha, K. (2004) Two-factor structural model of determinants of Brazilian sovereign risk, Journal of Fixed Income, 14(2).

(61) Nelson, C., Siegel, A. (1987) Parsimonious modeling of yield curves, Journal of Business, 60, 4, 473-89.

(62) Pagès, H. (2001) Can liquidity risk be subsumed in credit risk? A case study from Brady bond prices, BIS Working Papers No 101.

(63) Pan, J. and K. Singleton (2008). Default and recovery implicit in the term structure of sovereign CDS spreads. Journal of Finance, 63 (5), p. 2345-2384.

(64) Pericoli, M., Taboga, M. (2006) Canonical term-structure models with observable factors and the dynamics of bond risk premiums, Temi di Discussioni, Banca D'Italia.

(65) Pinheiro, F., Almeida, C., Vicente, J. (2007) Um modelo de fatores latentes com variáveis macroeconômicas para a curva de cupom cambial, Brazilian Review of Finance, 5(1).

(66) Robert, C., Casella, G. (2005). Monte Carlo Statistical Methods. Springer Texts in Statistics.

(67) Rodrigues, J., Araújo, F., Moreira, M. (2000) Optimal Interest Rate Rules in In‡ation Targeting Frameworks, Working paper series 6, Central Bank of Brazil.

(68) Rudebusch, G., Wu, T. (2003) A macro-…nance model of the term structure, monetary policy and the economy, Working Paper, Federal Reserve Bank of San Francisco.

(69) Shiryaev, A. (1995). Probability, Springer Verlag.

(70) Spencer, P. (2008). Stochastic volatility in a macro-…nance model of the U.S. term structure of interest rates 1961–2004, Journal of Money, Credit and Banking, 40(6), 1177–1214.

(71) Spiegelhalter, D., Best, N., Carlin, B., van der Linde, A. (2002) Bayesian measures of model complexity and …t (with discussion). Journal of the Royal Statistical Society Series B, 64, 583–639.

(72) Schönbucher, P. (2003) Credit Derivatives Pricing Models, Wiley Finance.

(73) Shousha (2005). Estrutura a termo da taxa de juros e dinâmica macroeconômica no Brasil. Master's dissertation, PUC-Rio.

(74) Svensson, L. (1997) In‡ation forecast targeting: implementing and monitoring in‡ation targets. European Economic Review, 41, no. 6.

(75) Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. Journal of Financial Economics, 5: 177-188.

(76) Vicente, J., Tabak, B. (2008) Forecasting bond yields in the Brazilian …xed income market. International Journal of Forecasting, 24(3), 490–497.

(77) West, M., Harrison, J. (1997) Bayesian forecasting and dynamic models, 2nd ed., New York: Springer-Verlag.

(78) Wu, T. (2003) Stylized facts on nominal term structure and business cycles: an empirical VAR study, Applied Economics, 35(8), 901-906.