UNIVERSITÀ DEGLI STUDI DI SIENA

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA E STATISTICA

Giulio Tarditi

Affine Term Structure Constraints on Euribor data

n. 613 – Giugno 2011

Abstract - This article discusses some modifications to non arbitrage models described by an Ornstein-Uhlenbeck latent process and an affine dynamic system. The empirical analysis refers to Euribor rates, due to the leading role they have in financial markets, but also to help the replicability of the results due to their accessibility and gratuitous nature. The benchmark model belongs to the class of Affine Term Structure Models (ATSM), whom owe their popularity to the success of Duffie and Kan (1996). Nodes have been calculated recursively through the use of the Kalman filter, and hence have the corresponding bayesian interpretation. The proposals differ from traditional models on some constraints posed on certain model specifications that allow to identify different aspects of the term structure. Through a clear identification of the type of contribution that each factor can undertake, it is possible to define probabilistic structures with minimal residuals purified from the dominant systematic residues visible in classic model residuals. Term Structure properties seem to be be identified with greater precision, which in the authors opinion justifies the relaxation of the hypothesis due to the additional constraints. The empirical analysis tries to convey such findings, and reminds of possible evolving paths of this line of work, such as a different specification of the transition process or the relaxation of linear and gaussian nature.

JEL Classification: C32, E43, G12, O52

Giulio Tarditi, Dipartimento di Economia Politica e Statistica, Università di Siena - giuliotarditi@yahoo.it

1 Introduction

The necessity to extract information from the ever growing amount of financial data has generated one of the most prosperous literatures within econometrics. New tools have been developed, some objectives have been reached, but difficulties persist whom require the evolution of models to higher levels of detail and complexity, in order to allow the comprehension of links between different interest rates, hence confirm or exclude the behaviour of agents in this sector. Interest rates are crucially important as they are often considered a barometer of the correct functioning of an economy, as they describe the intertemporal evaluation of investments through the monetization of the presents opportunity cost with respect to the future. Hence it is unrefusable that interest rates have an impact on the economy globally and indiscriminately.

The study of interest rates consists in the longitudinal analysis of various time-series, each determined by a specific maturity date. If the time-series were unrelated, the best instrument to analyse the variation of interest rates would probably be some model similar to aVAR[1], however this is clearly unrealistic as it would imply that the variation of a three month interest rate would not be taken into the economic decision making of financial agents concerned with a four month interest rate. The VAR model does not incorporate *a priori* information that is a known component of the logical framework, and the theory necessary to analyse interest rates is concerned and has as primary objective to correctly model such information. The correct probabilistic model is hence the one that is most restrictive, hence that excludes results incongruous with the theory, but is simultaneously sufficiently flexible to recognize all the complexities of a term structure, adapting to a variety of empirical observations. To understand which solutions must be excluded it is mandatory to comprehend why financial markets may be seen as efficient, and how this may exclude profits in absence of risk. The exclusion of these solutions defines non-arbitrage models[2], that integrate the notion of efficient markets[3].

Interest rates however are not only constrained to market efficiency, but also have a plausible interval of reference, which defines extreme values that are incompatible with the functioning of the economic system as we know it to be. The limits of this interval are unclear, and a subject of research in their own right,

[1] *Vector Auto Regression* (see [30]).

[2] see [5]

[3] see [13]

as periods of hyperinflation have reported very high interest rates due to the fact that purchasing power is the leading force behind utility maximization. However, these periods were highly dysfunctional, as double digits interest rates strain the economic functioning in various ways, hence do not last for long. What is generally recognized is that there are "common interests", promoted *in primis* by central banks, that require a stable interest rate and possibly an optimal long term value. This is the conceptual basis that justifies the hypothesis that the interest rate is a process subject to *mean reversion*. This means that although the process is disrupted by *shocks* of various nature, it is controlled by powers that restore its value to the optimal one for the economic system. A model that does not account for this additional information, would not exploit the *a priori* information, and consequently would not allow a credible probabilistic structure as one reaches values where these forces act, setting positive probabilities to the impossible states.

As it will be shown, the hypothesis underlying the classical model that describe the *a priori* properties of the economic system have been tested[4] as is indicated by the greater than average residuals during the last period. The information that we can extract from the past will necessarily be predictive of future economic developments, and can be useful to inform operators or identify in advance economic malfunctioning. The empirical analysis explored in the thesis of which this article presents the main results[5], compares different ATSM and evaluates the compromise between the adaptation to the available data and the complexity of the model. First varying the number of latent factors, which has not been here reported, and in second place through the proposal of original measures to identify the known different components of the term structure such as the level the gradient and the curvature.

The connection between the flexibility of the model to the empirical data will give a measure of the credibility of the working hypothesis and the role of the variations of the interest rates. In these models we assume that a known number of latent factors evolves through a specified relationship (transition equation) contaminated by gaussian disturbances. From such latent factor we can determine a new functional relationship that will provide estimates of the observed factor which are then compared to the observations. In other words we are estimating the signal in the data, which has been successfully developed in engineering under the name of Optimal Filtering.[6] However, as we have said interest rates require the

[4] especially during the financial crisis when liquidity was not guaranteed

[5] see [31]

[6] see [18]

incorporation of all the limitations expressed by financial mathematics, excluding the possibility of arbitrage within the process. This determines a relationship between the term structure and the latent variable. Through such hypothesis we are enriching the informative context regarding the nature of the underlying process, introducing the notion of efficient financial markets, as the necessary hypothesis[7] seem plausible at least for a first approximation of the phenomenon. ATSM are an extremely flexible class of models, subject to restrictive working hypothesis. The parametric space allows us to capture nearly all the variation of the evolution of the term structure, which is a critical requirement in order to appear realistic and consequently able to simulate alternative scenarios. Parameters tend to have a clear mathematical interpretation, which often becomes a clear economic interpretation. It allows us to incorporate or test economic theories, such as a long term mean attraction point or the limited interval of variation. Multifactor extensions extend the number of latent variables, hence allowing for a greater degree of uncertainty that guides the evolution of the data, seem substantially more realistic. The mathematics are slightly more complicated, but conceptually the extension is straightforward and intuitive, and suggests important guidelines for future developments.

[7] hypothesis such as negligible transaction costs or perfect information

2 The underlying logic

When we refer to a specific statistical model, we refer to a set of rules that specify the probabilistic context. The observations can be seen as the input from which we estimate the parameters that are most likely to have determined such data, given the probabilistic structure. This can be obtained through the minimization of a loss function that weights appropriately the discrepancies between the theoretical model and the dataset. What follows is a brief description of the working assumptions necessary to the class of ATSM.

Graphical interpretation: the markovian process

The first step necessary in the definition of ATSM is the specification of the dynamics of the term structure through one or more factors. These generic factors reflect the underlying uncertainty that constructs the yield curve, hence the term structure. The factors can be deterministic (known functions of time), adaptive processes (as GDP or CPI) or latent variables. In this analysis we will limit the scope to latent variables, excluding the interaction of other variables which can be later introduced[8]. The independence relations between these variables can be described by a diagram, in which we postulate the relation of causality by limiting the analysis to Direct Acyclical Graphs (DAGs). The simplest one factor model can be described by a Hidden Markov Model (HMM) with a single uncertainty factor, from which spawns at each temporal instant a single node of observed variables[9]. This node has the dimension of the number of observed time-series, and the latent factor can be extended to multiple dimensions with vector notation. Once we have specified the independence structure, such as to mirror the the complexity of the latent structure that generated the data, it is necessary to specify the infinitesimal dynamics of transition which determine how the data evolves through time. The terms “state”, “node” and “variable” are used interchangeably.

Local dynamics: the latent variable

Within the family of ATSM there are many possible specifications, many of which have as defining characteristic the formulation of local dynamics postulated over the latent variable that defines the interest rate. These variables are often identified with the notation $\{x(t)\}$.

[8] see [2]
[9] see figure 1

The dynamics can be expressed as Stochastic Differential Equations (SDE) whom describe the asymptotic and conditional probability distribution, hence the probabilistic path. The basic idea is extendable to independent multifactor models, and all that is needed to take into account multidimensional latent factor spaces is a change in notation. The fundamental working hypothesis can be expressed as:

1. The stochastic process of the interest rate is markovian[10].

2. The interest rate is guided by a latent variable through a known function.

3. The latent variable process is describe by a SDE as follows:

$$dx(t) = \underbrace{\alpha dt}_{drift} + \underbrace{\beta dW(t)}_{diffusion} \tag{1}$$

In this case $\{\alpha, \beta\}$ are generic functions $\{\alpha(r(t),t), \beta(r(t),t)\}$ and $dW(t)$ is the first derivative in t of a Wiener process defined on a generic probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The relationship between these two infinitesimal quantities identifies the functional form of $\{\alpha, \beta\}$ and how such components influence the drift and the diffusion of the stochastic process.

Through such specification we can evaluate an expected value and a variance which can be expressed by parameters. We can choose functional forms that simplify calculations and that are able to bind the probabilistic structure to properties derived from economic theory. We can guarantee that the interest rate will have an attractor point, whom will maintain the evolution of the process within credible intervals and that does not allow impossible realisations. One important property is a limited asymptotic variation, as in real life there are powers that monitor and guarantee a the functioning of the economic system through monetary policies.

Affine transformation: the observed variable

As we try to model the term structure, we limit our analysis to models whom are affine in nature[11] due to the important computational simplifications. Such limitation allows us to obtain an explicit solution, and hence focus on the issue of estimating correctly the parameters. When dealing with models as large as twelve

[10] This is justified by the efficient market hypothesis, which states that all past information is irrelevant to future forecasting given the present state

[11] Affine Transformation: $\mathfrak{F} : \Re^n \to \Re : \exists \left\{ f(\mathbf{x}) = a + \mathbf{b}^\top \mathbf{x} \right\} \forall \mathbf{x} \in \Re^n; \forall a \in \Re; \forall \mathbf{b} \in \Re^n$

dimensions likelihood computation becomes increasingly complex, and may require Markov-Chain Monte-Carlo (MCMC) tools that however are not discussed here.

ATSM are completely specified by the instantaneous interest rate. This is possible through no-arbitrage reasoning, which is in some way similar to the derivation of the Black-Scholes equation, as it obtains a Partial Differential Equation (PDE) from a SDE through the use of Ito's lemma. The PDE solution expresses the function that determines the price of Zero Coupon Bond, hence links the latent variable to the term structure.

ATSMs are non-arbitrage models such that the *spot* interest rate is an affine function of the vector of latent variables[12]:

$$r_t^{\tau} = a(\tau) + \mathbf{b}(\tau)^{\top} x_t \tag{2}$$

$$\mathbf{r}_t = \mathbf{a} + \mathbf{B}^{\top} \mathbf{x}_t \tag{3}$$

Data interpretation

However the latent variable is not directly visible from data, i.e. it is not empirically recorded, hence we must compare the data with the observed variable $\{y(t)\}$ and link such variable to $\{x(t)\}$ through a specific functional form. This functional form is the crucial node through which we introduce all the non-arbitrage hypothesis, and trough which we filter data through a scheme that takes into account the impossibility of solutions excluded by non-arbitrage boundaries.

We must extract the information pertinent to latent variables from the realization of observed variables, a field of research intensely developed in a very different context, where algorithms have been studied in order to optimize the extraction of information sent through a disturbed channel. The use of the Kalman Filter (KF)[13] to estimate the expected path of latent variables allows the use of state space notation, under which gaussian simplifying assumptions guarantee important asymptotic results and explicit solutions.

From a statistical point of view, the use of the KF is of great importance as it corresponds to a specific recursive bayesian estimator, hence it provvides a complete description of prior, likelihood and posterior distribution from which we can conduct a bayesian analysis. From a theoretical stand the gaussian framework

[12] $\tau = T - t$

[13] see [4]

gives optimal estimates from more then one point of view. The updated observed variable estimates are *Maximum A Posteriori* (MAP) estimates and also *Maximum Mean Square Error* (MMSE) as we are dealing with a symmetric distribution.

This model hopes to shed light on the first steps towards the development of models that are able to relax gaussianity assumptions through the substitution of the KF with the *Particle Filter*[14] (PF). However the PF does not have the elegance of an explicit solution, and needs Monte Carlo (MC) simulations which are computationally demanding and aggraviate parameter estimation difficulties. Through the use of market efficiency to specify the transition equation, and non-arbitrage theoretical fundamentals to define the measurement equation, all that is left to estimate the parameters is to arbitrary initialize the maximization of the likelihood function.

3 The classic model

The exclusion of arbitrage possibilities justifies the use of a PDE in substitution of the SDE that defines the evolution of the latent variable, by constructing a deterministic portfolio, hence risk-less, obtained through the investment of bonds with different maturities.

A risk-less portfolio will necessarily return a risk-less interest rate, to avoid risk-less profit strategies. Returns above the risk-less rate must then be justified by a constant value that relates the return proportionally to the amount of risk undergone. This constant is the market price of risk, and can be interpreted as the Radon-Nikodym derivative that allows the change in measure from the real measure to the risk neutral measure.

The instantaneous spot rate can be specified by the following SDE:

$$d\mathbf{x}_t = a(\mathbf{x},t)dt + b(\mathbf{x},t)d\mathbf{W}_t \tag{4}$$

Applying Ito's lemma we obtain the following results:

$$dP(t,T,r_t) = [P_t + \frac{1}{2}P_{rr}b(\mathbf{x},t)^2 + P_r a(\mathbf{x},t)]dt + [b(\mathbf{x},t)P_r]dW_t \tag{5}$$

This is a SDE with drift parameter $P_t + \frac{1}{2}P_{rr}b(\mathbf{x},t)^2 + P_r + a(\mathbf{x},t)$ and diffusion coefficient $b(\mathbf{x},t)P_r$. The revolutionary contribution of Vasicek transforms this SDE into a PDE that is independent of the stochastic component expressed by the

[14] see [8]

Wiener process. Similarly to the Black-Scholes model, one can have a portfolio of bonds able to replicate the return of a deterministic portfolio.

The deterministic portfolio

Let $\{V\}$ be the value of a bond portfolio whose underlying assets have maturities and current prices as follows $\{T_1;T_2\}$ $\{P_1(t,T_1);P_2(t,T_2)\}$. If the weights that specify the composition are $\{u_1,u_2\}$. The instantaneous return is then expressed as:

$$\frac{dV_t}{dt} = u_1\frac{dP_1(t,T_1)}{P_1(t,T_1)} + u_2\frac{dP_2(t,T_2)}{P_2(t,T_2)} \tag{6}$$

By selecting appropriate weights we can annihilate the stochastic component, and once we substitute such weights in the above equation, we obtain the following risk-less portfolio[15][16]:

$$\frac{dV_t}{dt} = (\mu_2\frac{\sigma_1}{\sigma_1-\sigma_2} - \mu_1\frac{\sigma_2}{\sigma_1-\sigma_2})dt \tag{7}$$

Through non-arbitrage logic and mathematical substitutions we can obtain the following equation that links the market price of risk[17] to the instantaneous drift, diffusion coefficient and risk-less interest rate:

$$\mu - r_t = \sigma\lambda_t \tag{8}$$

This notation highlights how the market price of risk is proportional to the return in excess of the risk free rate, for all underlying maturities, and the multiplicative constant is given by the measure of risk. Similarities with the Capital Asset Pricing Model are noticeable[18], where the expected return is linked to the risk of a share.

[15] $\mu_i = \frac{P_{i,t}+\frac{1}{2}P_{i,rr}b(\mathbf{x},t)^2+P_{i,r}a(\mathbf{x},t)}{P_i(t,T_i)}$ $\forall i \in [1,2]$

[16] $\sigma_i = \frac{b(\mathbf{x},t)P_{i,r}}{P_i(t,T_i)}$ $\forall i \in [1,2]$

[17] $\lambda_t \triangleq \frac{(\mu_1-r_t)}{\sigma_1} = \frac{(\mu_2-r_t)}{\sigma_2} = \frac{(\mu-r_t)}{\sigma}$

[18] see [25]

The affine model

In this analysis we are concerned exclusively with ATSM, and in particular with a generalization of the Vasicek model[19], although it could be extended to the alternative Cox, Ingersol and Ross (CIR) model[20], hence we will proceed with a generic specification and introduce the Vasicek specification later. ATSM rely on restrictive hypothesis, that are however increasingly flexible as we allow the number of factors to grow. The class of affine model includes all models that can be expressed as follows:

$$P(\tau) = e^{A(\tau)-B(\tau)r} \tag{9}$$

We can identify many different specifications, many of which share various desirable analytical properties, that simplify the solution of the PDE and often allow an explicit solution. This solution will link the state variable to the price of the asset, in a way that the term structure is completely specified by the interest rates through the exclusive use of the present state variable.

Having specified equation(9) it is now possible to compute the necessary partial derivatives to resolve the PDE:

$$(-A'(\tau)+B'(\tau)r)P(\tau)-(a(\mathbf{x},t)-b(\mathbf{x},t)\lambda_t)B(\tau)P(\tau)+\frac{b(\mathbf{x},t)^2}{2}B^2(\tau)P(\tau)-r_tP(\tau)\doteq 0 \tag{10}$$

$$-A'(\tau)-(1-B'(\tau))r-(a(\mathbf{x},t)-b(\mathbf{x},t)\lambda_t)B(\tau)+\frac{b(\mathbf{x},t)^2}{2}B^2(\tau)-r_t=0 \tag{11}$$

From which we can obtain through the introduction of generic affine specifications[21]:

$$\left(-A'(\tau)-\alpha_1B(\tau)+\frac{1}{2}\beta_1B^2(\tau)\right)-\left(1-B'(\tau)+\alpha_0B(\tau)-\frac{1}{2}\beta_0B^2(\tau)\right)r_t=0 \tag{12}$$

From (12) we can verify that the only non contradictory solution for all values of r_t and τ is the case in which the following system of ODE is coherent:

---

[19] see [32]
[20] see [9]
[21] $\tilde{\mu} \triangleq (a(\mathbf{x},t)-b(\mathbf{x},t)\lambda_t)=\alpha_0r+\alpha_1$; $\tilde{\sigma} \triangleq b(\mathbf{x},t)=\sqrt{\beta_0r+\beta_1}$

$$-A'(\tau) - \alpha_1 B(\tau) + \frac{1}{2}\beta_1 B^2(\tau) \doteq 0 \tag{13}$$

$$-B'(\tau) + \alpha_0 B(\tau) - \frac{1}{2}\beta_0 B^2(\tau) \doteq -1 \tag{14}$$

We must also introduce the boundary value, as we know that the underlying argument is a ZCB and that the value at maturity is by definition unitary.

$$P(T,T) = e^{A(0)-B(0)r} = 1 \Rightarrow A(0) = B(0) = 0 \tag{15}$$

The Vasicek model

At this point it is convenient to specify the explicit functional form of $\tilde{\mu}$ and $\tilde{\sigma}$. Following the Ornstein-Uhlenbeck specification:

$$\tilde{\mu} \triangleq k\left(\overline{\theta} - r_t\right); \tilde{\sigma} \triangleq \sigma \tag{16}$$

$$\overline{\theta} = \theta - \frac{\sigma\lambda}{k} \tag{17}$$

Such that the underlying process is defined as:

$$dr_t = k(\overline{\theta} - r)dt + \sigma dW_t \tag{18}$$

In this case, after some tedious mathematics, the term structure is completely defined as:

$$P(\tau, r) = \exp^{A(\tau)-B(\tau)r} \tag{19}$$

$$A(\tau) = \frac{\gamma(B(\tau) - \tau)}{k^2} - \frac{\sigma^2 B^2(\tau)}{4k} \tag{20}$$

$$B(\tau) = \frac{1}{k}\left(1 - \exp^{-k\tau}\right) \tag{21}$$

$$\gamma \triangleq k^2\left(\theta - \frac{\sigma\lambda}{k}\right) - \frac{\sigma^2}{2} \tag{22}$$

The Vasicek model is a model in which the instantaneous interest rate follows the continuous version of the autoregressive process. This link can be shown by using an alternative notation for which $\triangle t = 1$ identifies an autoregressive model.

$$\triangle r_t = k\left(\theta - r_{t-1}\right)\triangle t + \sigma \triangle z_t \tag{23}$$

$$r_t = k\theta - \left(1-k\right)r_{t-1} + \varepsilon_t \tag{24}$$

The time interval need not be discrete as in the difference notation above. We can interpret the equation in continuous terms by letting the time interval tend to an infinitesimal amount. Under the notation or which $\triangle z_t \sim N(0, \triangle t)$ we have the following already mentioned equation:

$$dr_t = k\left(\theta - r_t\right)dt + \sigma dW_t \tag{25}$$

The θ parameter identifies the value to which the process converges in the long term, k identifies the speed of convergence to such value and σ is a measure of the instantaneous volatility. The only parameter not specified in the transition equation is the market price of risk λ, as it models the link across maturities.

Similarly to the autoregressive process, $k = 0$ identifies a *random walk* process which is memoryless and does not converge, low values identify strong persistence in the process, and high values a rapid return to the asymptotic distribution. The main limitations of this model are given by the fact that it does not exclude negative interest rates and volatility is independent of the value of the interest rate.

4 The original contribution

Additivity

The main contribution is in the specification of some constrained models that take into account previous thoughts and considerations, and attempt to correct certain aspects which are not treated in the traditional model. In particular:

1. A latent factor is most useful if it identifies a certain aspect of reality, hence it is a synthetic measure of a certain type of information. In order to differentiate and specialize they must be defined in different ways. Constrained factor models contribute with an original specification of such definition.

2. Affine models are fundamental blocks to many successive models, however the affine characteristic is inherently a limitation to the identification of non linear characteristics such as the curvature. The models here presented allow a partial relaxation of this hypothesis, in order to identify the third factor.

Additivity -through factors- in affine models is one of the most intuitive and useful properties. It allows the decomposition of any observed factor in the sum what we will call "contributions", hence we can observe it at any time t as the sum of individual parts. These contributions can be either a constant benchmark or the product between a latent factor and the corresponding multiplicative systematic component. The additive component can be seen as the constant contribute over which each factor manifests its additive effects. There could be a further decomposition of the additive effect, although it has no macroscopic meaning of interest.

The latent factor has a direct impact on the corresponding contribution. If we artificially half the first latent factor it will *ceteris paribus* half the effect of the first contribution on the observed factor. Clearly the same is true for the multiplicative component. We can then invert the reasoning and consider the argument of main interest the contribution, not the latent factor or the multiplicative component, as it is the sum of the contributes that constructs the observed factor, and consequently directs the choice of parameters depending on the discrepancies with the observations.

If we multiply the multiplicative component by a some value and the latent factor by it's inverse, the effects will cancel out and we will obtain the same residuals. We must then comprehend that the scale of reference of a factor can be counterbalanced by a flat multiplicative component. The optimal choice of parameters is composed of two parts, the discrepancies between the observations and sum of the contributions and the specification of the properties of the theoretical process. The first is the argument of evaluation and the second is the method of evaluation. The constrained models want to artificially separate these parts, in order to better comprehend these aspects by constraining each factor to a single effect form a contribution stand, even if it is left free from a distribution property point of view.

The idea proposed to constrain the multiplicative components is basic. In the two factor model the multiplicative component of the second factor is constrained by translation to have mean zero. With such procedure any choice of parameters will uniquely specify “zero mean effects”. In the three factor model the curvature is accentuated through a second degree function compensated for level and slope. The third factor will then have a comparative advantage to describe the curvature in the term structure. These modifications not only clarify the identity and the importance of the latent factors, but also yield lower residuals which means that the structures seem to be significantly present in the data generating structure. This new specification allows the latent factors to keep the same transition equation, hence have the same asymptotic and conditioned properties. The first fifty days are reported[22], and visualise the decomposition of the observed factor in first contribution, second contribution, third contribution, additive contribution (or component), observed factor and observation.

One can also see the evolution of the four contributions[23] (respectively additive, level, slope and curvature) on the whole period of study, while underneath the same information is depicted deprived of the temporal dimension in order to evaluate the magnitude and interval of variation of each contribution. The color gradation allows to identify the temporal dimension more clearly, while in the corresponding collapsed graph it is in rotation through time, hence does not identify any particularly useful information.

Latent factors

Having clarified the aim of the constrained models we report the results obtained. A complete comparative analysis of the three traditional models (one, two and three factors) with the two constrained models (two and three factors) is fully disclosed in the original thesis. In such analysis the interest rate “level” identification is evident in all models, but while in the traditional models the remaining variation is divided through the factors, in the constrained models the factor specificity

[22] see figure 5

[23] see figure 6

allows each one to identify the optimal parameters that describe each type of variation. The effects of such a modification on the non-arbitrage assumptions are not trivial, as the contribution of the second factor adapts to a second sub-problem unresolved by the first factor, in which the original assumptions may not apply. Hence we tend to evaluate the interpretation of the results and the minimization of the residual discrepancies over the importance of the underlying assumptions of which these models are an extension and possibly a violation. The factors can represent a measure of the underlying causes that motivate financial agents decision making, which can themselves have a trend or represent a process with intrinsic scientific interest.

Residuals

Residuals in constrained models are significantly reduced with respect to previous models[24]. This reduction is less evident once one incorporates present observation information, i.e. once the latent factor is updated with the present information, however the minimization of residuals once present information is available has clearly a different interpretation, and has not been pursued. In the two factor model residuals are small and zero-centered[25]. This reduction could have a negative impact on the thick tail problem, typical of financial data, by endangering the gaussian hypothesis. However in this case it is of no major concern, as the residuals are zero-centered and the error persistence in the one factor model has been eliminated[26]. The mean residual across maturities is not correlated with the adjacent time intervals. However, as the third factor has not been introduced yet, one can notice that the minimum and maximum interest rates maturity tend to simultaneously have lower (o higher) values then the intermediate maturities, which would suggest a certain curvature of the residuals across maturities.

In the constrained three factor model the residuals have substantially improved with respect to all other models[27]. The residuals are smaller and one cannot identify any systematic pattern. The third pattern -responsible for the change in curvature- seems to be sufficient to describe all other aspects that could have been covered by other additional factors, such as change in symmetry or bimodal structures. The variability is proportional to maturity except for particularly informative time periods during which the short term rate shows unusual variation. The

[24] most fitting traditional model in figure 2
[25] see figure 3
[26] here not shown
[27] see figure 4

residual error is not constant through time, as specified in the model, which was foreseeable as various monetary policies have been implemented during the time frame. An extention could model dynamical parameters in order to measure the dynamics of a shock on the monetary system.

The fit of a model is evaluated by weighing the residuals with respect to the generating probabilistic structure, hence lower residuals contribute *caeteris paribus* to support the evidence in favour of a certain model. Of equal importance to a correct specification of a model is the complete extraction of systematic variation from the observations and hence residuals with no recognizable structure. Probabilistically linked residuals through a systematic distortion clearly violate one or more working hypothesis.

However in this specific case the evidence of the model not perfectly adapting to the model in the last time frame is actually a element of support to the model, as the time interval coincides with the housing bubble burst and the financial crisis which is recognized as an anomaly, during which markets were no longer efficient due to liquidity shortage in inter-banking loans. The model in this case is an indicator of the system's correct functioning, hence the heavy tailed residuals are of more interest then coherent data that does not show or take into account this anomaly.

Another particularly interesting time interval is given by the first fifty days from the birth of the Euribor. One can immagine that the creation of a complicated system as the inter-banking market on a common european currency -which did not physically exist yet- would not perfectly adapt to a theoretical model. In these days the yield curve appears inverted, which is rare as the interest rate on long maturities are lower then on short maturities. Due to the uniqueness of this period, which presents extreme characteristics that can be modeled with difficulty, a comparison between the five different models has been focused on this particular aspect. The single factor model shows all its limits, not being flexible enough to incorporate the information. Two and three factor models also find difficulties on this time frame. On the other hand, the flexibility of the "constrained" models allows to adapt to the unusual nature of the data separating such occurrence from temporary, fluctuating, short lived shocks, as one would expect from a correctly specified model.

Likelihoods

Discussing likelihoods in traditional models presents numerous difficulties, due to the possible factor interactions and compensations, and the unclear effects of parameter variation. Contrary to traditional models, constrained models generally have a simple interpretation through the recognition of the contribution to the model. It is necessary to develop an in-depth study, on the parametr space and how the likelihood varies through time, to identify the representation of a specified factor by the choice of parameters, but also the choice of parameters that identify a similar process. Then one could establish a clear parameter identification that takes into account the implications on asymptotic and conditional properties of the generating probabilistic process. Here we present a relatively convincing choice of parameters, knowing that they are not conclusive and are only the tip of the iceberg. It is also important to highlight that constrained models perform better then traditional models not due to an ad hoc choice of parameters, but because the increased flexibility caused by the relaxation of the underlying hypothesis, allows to recognize the three distinct aspects of the dataset. To each of these aspects correspond distinct economic behaviour. Parameters have been analysed by observing contour-likelihoods in key parameter sets, however this is left as a suggested future development. In tables 1 and 2 we report the chosen parameters to allow a replication of the results, and a verification of likelihood calculations or squared discrepancies.

5 Conclusion

The comparison between the different models has animated many thoughts on the nature of the data, but the author has chosen to give a coherent practical description instead of an *a-priori* theory on the properties that markets should have. Regarding hypothesis coherence, graphic and a numeric measures are computed throughout the time frame.

The traditional model with a single factor has appeared strongly inadequate, mainly due to the systematic component left in the residuals. New uncertainty sources are introduced with additional latent factors, which have immediately showed a significant improvement, by reducing and in part eliminating the systematic component in the residuals. By allowing to synthesize the information of the yield curve in two distinct values, these move inversely in such a way to produce a combined optimal description of the yield curve. In the traditional model the third factor does not contribute substantially to the residual reduction, and the meaning or utility of this new factor is not evident. The parameter overlap make a unique identification of the process components very difficult, difficulty that jointly with the parameter space search casts serious doubts on the effective optimality of the results obtained.

The addition of constraints on the multiplicative components of the model restrain the contribution of each latent factor to a identifiable characteristic type of the dataset. The factors are then able to find the uncertainty proportion of the process that most reflects their comparative advantage on the others, hence contributing in different ways on different parametric spaces. As the model is additive each contribution is separable from the resulting observed factor, and can be used as a measurement of the behaviour or preferences of the market agent. The specification of the contribution is then also a useful tool to steer optimal parameter search. The optimality of the additive and multiplicative components on the residuals -as discrepancies from the observations- must however be distinguished from the probabilistic optimality of the underlying process -as asymptotic and conditional process properties.

There is a clear relationship between the non-arbitrage hypothesis constrains that limit the flexibility of what can be modeled, and the complexity of strategic behaviour within financial markets. The affine structure allows to identify two different aspects: the first is given by the specification of the additive and multiplicative components, the second is given by the probabilistic properties of the underlying process described by the optimal parameter set. This decomposition is presented as a useful tool for the understanding of interest rate interaction and yield curve dynamics.

Concluding, through the introduction of constrained factors one can allow extra flexibility to the model, however this implies a cost in the relationship between the original processes and the model hypothesis. A radical specification of the multiplicative component allows to constrain the contributions and better comprehend the dynamic evolution of the cross-section properties such as level, slope and curvature.

6 Appendix

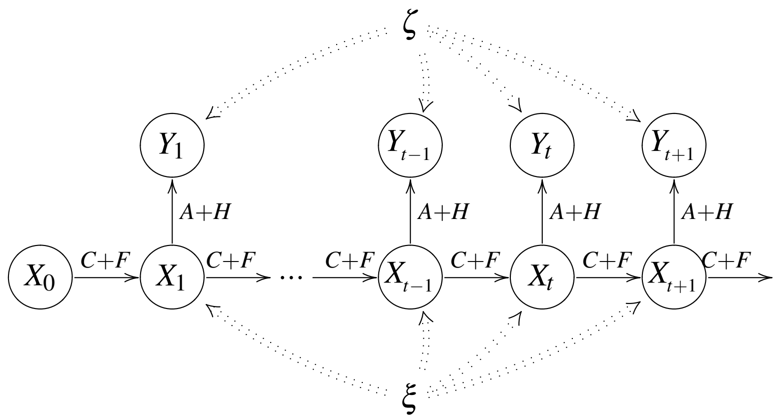

Fig. 1: Space state in the linear gaussian case

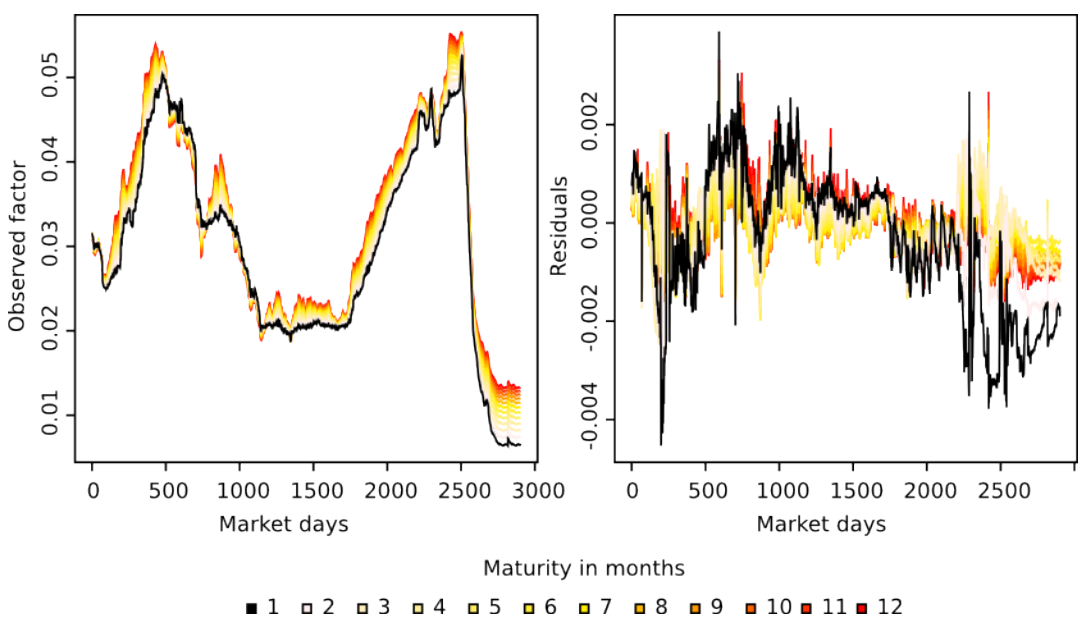

Fig. 2: Traditional three-factor model residuals

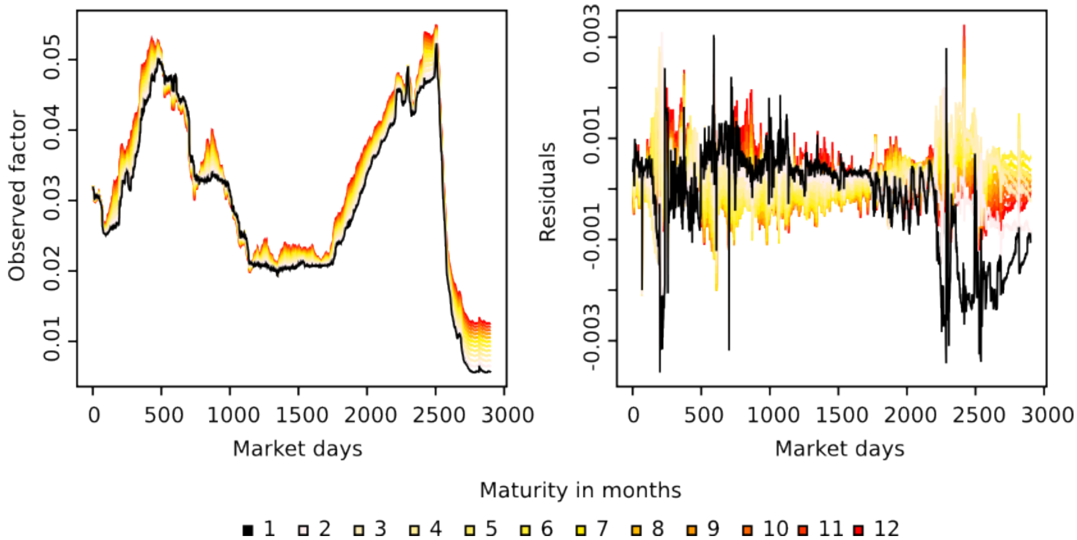

Fig. 3: Constrained two-factor model residuals

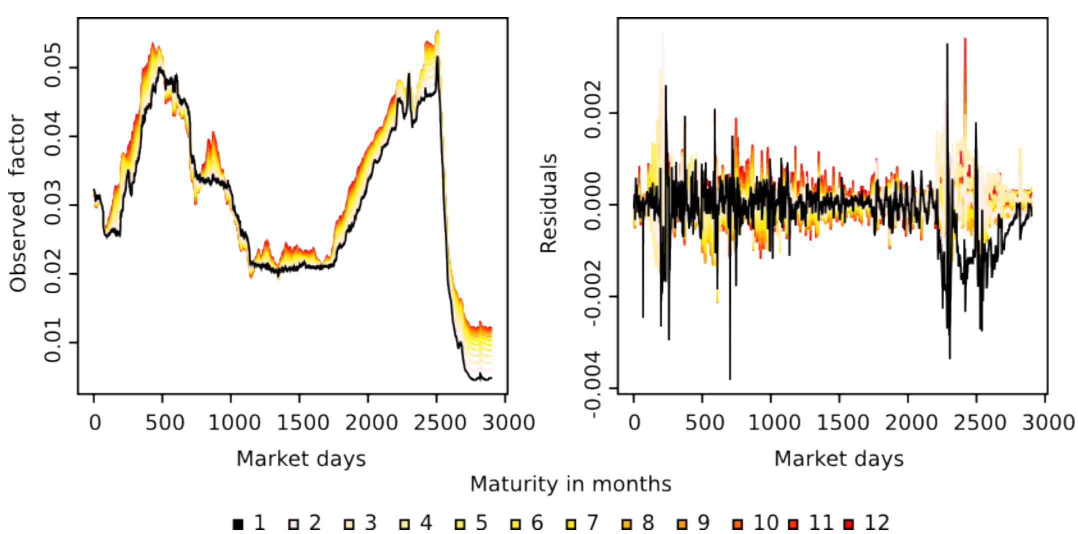

Fig. 4: Constrained three-factor model residuals

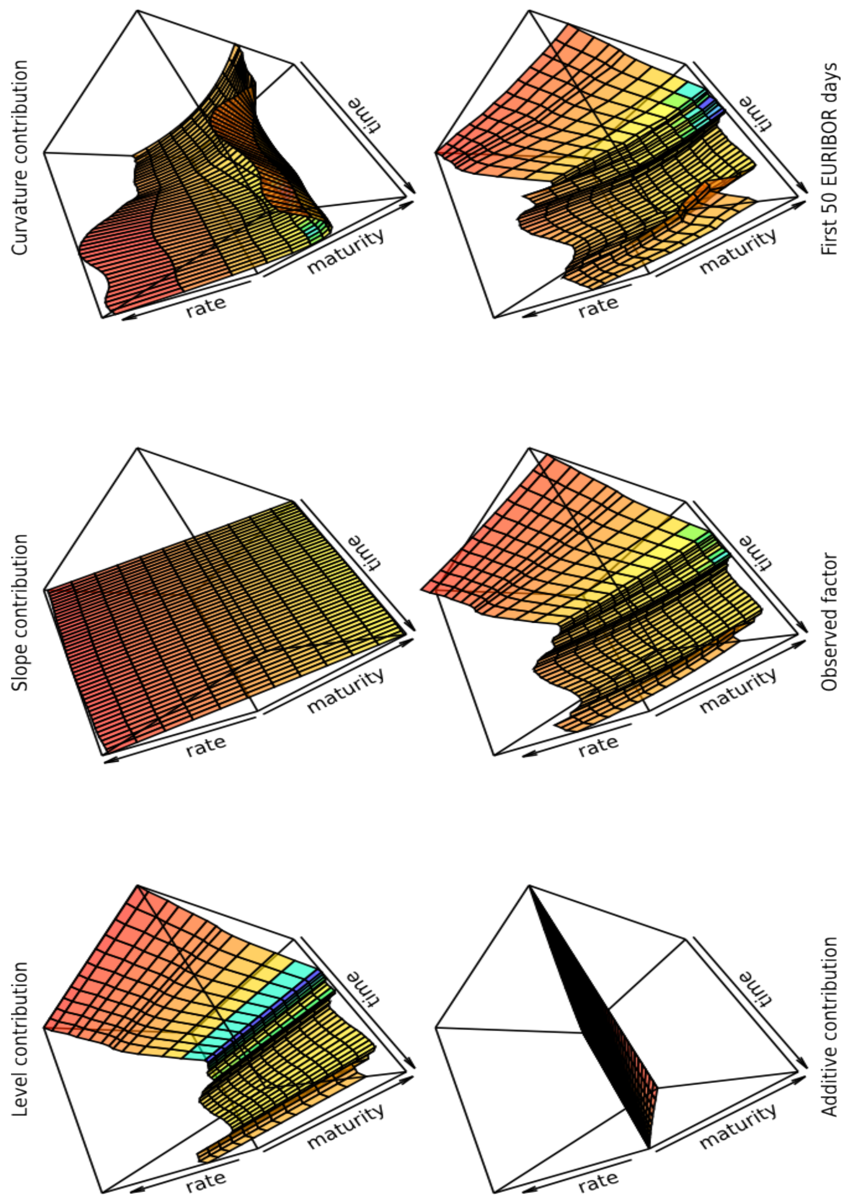

Fig. 5: Observed factor decomposition

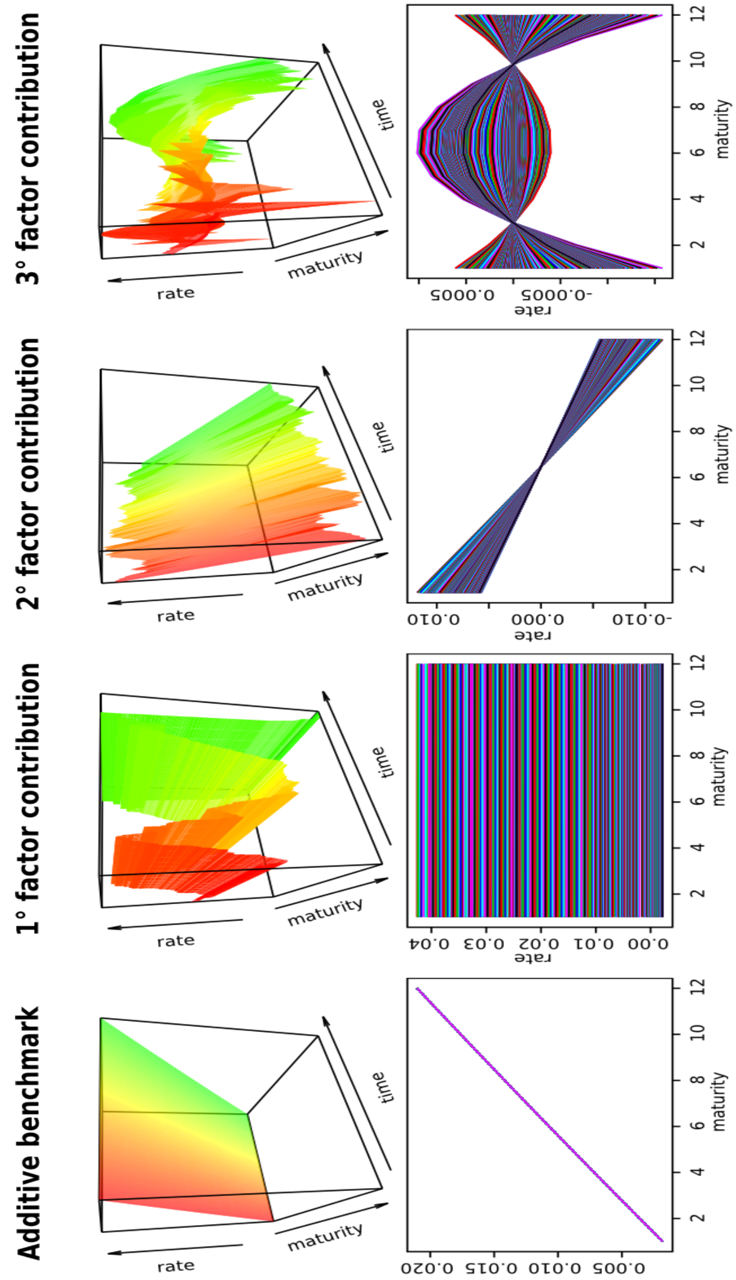

Fig. 6: Contribution variation through time

Parameters	1^{st} factor	2^{nd} factor
k	9.237048e-02	1.088308290
θ	6.799117e-02	-2.382250e-01
σ	-4.105327e-03	-5.178111e-03
λ	-5.014241e-02	2.085541110
$\propto -\ell(x\|k,\theta,\sigma,\lambda)$	-232659.0	
$\sum_{t=1}^{T}(y_t-\hat{y}_t)^2$	0.009489443	

Tab. 1: Optimized parameters: two factors

Parameters	1^{st} factor	2^{nd} factor	3^{rd} facto
k	4.632832e-04	9.427588e-02	1.682986e-01
θ	-2.503774e-01	3.813946e-01	-1.884196e-03
σ	3.948297e-03	-1.844981e-02	9.408281e-04
λ	-2.569966e-02	3.285493e-01	-1.9680491752
$\propto -\ell(x\|k,\theta,\sigma,\lambda)$	-234550.5		
$\sum_{t=1}^{T}(y_t-\hat{y}_t)^2$	0.005028747		

Tab. 2: Optimized parameters: three factors

References

[1] Ait-Sahalia, Y. and Hansen, L. (eds.). (2009). *Handbook of Financial Econometrics* 1.1. ELSEVIER, Amsterdam.

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

[3] Arrow, K. J. and Fisher, A. C. (1974). Environmental Preservation, Uncertainty, and Irreversibility. *Quarterly Journal of Economics*, **88** (1), 312-19.

[4] Babbs, S. H. and Nowman, B. K. (1999). Kalman Filtering of Generalized Vasicek Term Structure Models. *The Journal of Financial and Quantitative Analysis*, **34** (1), 115-130.

[5] Björk, T. (1998). *Arbitrage Theory in Continuous Time*. Oxford University Press, Oxford.

[6] Bliss, R.R. (1997). Movements in the Term Structure of Interest Rates. *Federal Reserve Bank of Atlanta Economic Review*, 16-33.

[7] Bolder, D. and Streliski, D. (1999). *Yield Curve Modelling at the Bank of Canada*. Technical Report No. 84. Bank of Canada, Ottawa.

[8] Chen, Z. *Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond*. Manuscript.

[9] Cox, J. C., Ingersoll, J. E. and Ross, S.A. (1985). An Intertemporal General Equilibrium Model of Asset Prices. *Econometrica*, **53**, 363-384.

[10] Culbertson, J. H. (1957). The term structure of interest rates. *Quarterly Journal of Economics*, **72**, 485-517.

[11] Dai, Q. and Singleton, K. J. (2000). Specification Analysis of Affine Term Structure Models. *Journal of Finance*, **55**(5).

[12] Duffie, D. and Kan, R. (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance,* **6**, 379-406.

[13] Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance,* **25** (2), 383-417.

[14] Fisher, I. (1896). Appreciation and Interest. *Publications of the American Economic Association*, Macmillan, New York.

[15] Fisher, I. (1930). *The theory of interest*. Macmillan, New York.

[16] Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

[17] Hicks, J.R. (1946). *Value and Capital*. Oxford University Press, London.

[18] Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering,* **82** (1), 35-45.

[19] Keynes, J.M. (1936). *The General Theory of Employment, Interest, and Money*. Harcourt, Brace & World, New York.

[20] LeRoy, S. F. and Werner, J. (2001). *Principles of Financial Economics*. Cambridge University Press, Cambridge.

[21] Litterman, R. and Scheinkman, J. (1991). Common Factors Affecting Bond Returns. *Journal of Fixed Income*, **1** , 54-61.

[22] Longstaff, F.A. and Schwartz, E. S. (1992). Interest Rate Volatility and Term Structure: A Two-Factor General Equilibrium Model. *Journal of Finance*, **47**(4), 1259-1282.

[23] Lutz, F.A. (1940). The Structure of Interest Rates. *Quarterly Journal of Economics*, **55**, 36-63.

[24] MacKenzie, D. (2008). What's in a Number?. *London Review of Books*, **30** (18).

[25] Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, **41**, 867-886

[26] Mises, L. von. ([1949] 1998). *Human Action*. Scholar's Edition. Ludwig von Mises Institute. Auburn, Ala.

[27] Modigliani, F. and Sutch, R. (1966). Innovations in interest rate policy. *American Economic Review*, **56**(2), 178-197.

[28] Mulligan, R.F. (2007). Property rights and time preferences. *The Quarterly Journal of Austrian Economics*, **10**(1), 23-49.

[29] Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, **29**, 315-335.

[30] Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, **48**, 1-48.

[31] Tarditi, G. (2010). Ph.D. Thesis. Università degli studi di Roma Tre.

[32] Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, **5**, 177-188.