A LEARNING HYPOTHESIS OF THE TERM STRUCTURE OF INTEREST RATES

BALÁZS ROMHÁNYI[1]

Abstract

Recent empirical results about the US term structure are difficult to reconcile with the classical hypothesis of rational expectations even if time-varying but stationary term premia are allowed for. A hypothesis of rational learning about the conditional variance of the log pricing kernel is put forward. In a simple, illustrative consumption-based asset pricing model the long-term interest rate turns out to have an economic meaning distinct from both price stability and full employment, namely to measure the market perception of aggregate level of future risk in the economy. Implications for economic modeling and monetary policy are explored.

JEL: D8, E43, E44, E52, E58, G12

Key words: term structure; interest rate; learning; uncertainty; monetary policy

1 Balázs Romhányi, (Hungarian Ministry of Finance) H-1025 Budapest, Csatárka út 42-50, Hungary; tel: (361)-327-5996; Fax: (361)-327-5656; e-mail: balazs.romhanyi@pm.gov.hu

1. INTRODUCTION

There is some inconsistency about the relation of macro variables and the term structure of interest rates in the science of economics and finance. The triangle of short term interest rates, long term interest rates and macroeconomic variables (mainly output and inflation) is always broken somewhere. Neoclassical macro models (e.g. IS-LM) talk about macro variables and interest rates, but implicitly they always mean medium and long term interest rates, because these are relevant for the actors of the real economy when they decide about consumption, investment and saving. Monetary models talk about macro variables and interest rate, but explicitly they always mean the short term interest rate, because this is the tool of the central bank to influence the economy. Finance models talk about short, medium and long term interest rates, but rarely talk about macro variables, because they are not useful enough for forecasting the movements in the yield curve or asset pricing on a daily basis. The only exception is the consumption based model of asset pricing, but it isn't really a macro model, since it doesn't feed back interest rates into the law of motion of macro variables and it isn't really a finance model, due to its poor practical performance (e.g. Duffie, 1996, p.229).

In the practice of finance there are several main families of term structure models (e.g. affine models, such as the Vasicek-model, market models such as the Heath-Jarrow-Morton model, or non-linear models), but none of them is linked either to preferences of market participants or production technology available in the economy. This is clearly in contrast with the theory of macroeconomics and the practice of monetary policy. Macroeconomic theory – according to the Fisher-hypothesis – states that risk free nominal interest rates are determined by expected rate of inflation and expected real interest rates, but expected real interest rates are linked at least on the long run to expected real return on real assets. The practice of monetary policy is that short term interest rates are almost completely controlled by central banks, which are free to decide, whether they follow some interest rate setting rule (e.g. the Taylor-rule) making explicit the quantitative effect of different macro variables on the evolution of short term interest rates or they pursue (at least explicitly) a discretionary monetary policy.

There is another inconsistency in theory concerning the role of financial markets in monetary policy. On one hand the standard macroeconomic approach postulates that market participants are rational and markets are efficient, so any information based on market prices are just reflections of expected developments in variables (output, inflation, central bank rate, etc), which are exogenous for the financial market. If financial markets function in an "efficient" way, then central banks might use market indicators as proxies for directly unobservable variables, which are exogenous for the central banks as well (e.g. supply side shocks). A long series of empirical papers put the question what sort of macroeconomic events might be forecast from changes in the yield curve (e.g. Stock and Watson 1989, Smets and Tsatsaronis 1997, Schich 2000, Hamilton and Kim 2000). On the other hand, a long strand of empirical research in financial economics and econometrics documents the different types of anomalies compared with possible implications of the standard theory. Some of them are treated as "irrationality", but some of them are theoretically justified, such as overshooting, Peso-problems, rational learning, or credibility problems, as economic policy experts label them.

The need for modeling financial markets in the frame of macroeconomic analysis is striking from the empirical results of Stock and Watson (2003). They calculate changes in volatility and cyclical correlations of major economic variables (GDP and its components, production, employment, inflation, short and long term interest rate) of the U.S. economy in the period 1960-2002. In the second half of the period (1984-2002) the standard deviation of all the real and nominal variables is significantly lower, than in the first half (1960-1983), except for one variable, which is long term interest rate (measured by the 10-year T-bond rate). Standard deviation of GDP, consumption and inflation (GDP deflator) fell by 39, 40 and 47 percent respectively, while the standard deviation of long term interest rate actually rose by 10 percent. The conclusion is rather strait forward: there is something missing in the list of macro variables from the key determinants of long term interest rates and the first candidate is some financial market indicator. Unfortunately not much has been done to find an indicator characterizing the stance of financial markets in the very way as the fed fund rate is the indicator for the stance of monetary policy [see Bernanke and Blinder(1992)].

A third inconsistency can be found in the way interest rates as stochastic processes are modeled in economics and in finance. There is just one economic theory of the term structure linking long term interest rates to short term interest rates, namely the expectations hypothesis (EH) first formalized by Hicks (1939) stating that long term interest rates are equal to the expected average of short term interest rates. Both the literature about refusing and saving the EH is already voluminous, but the evidence about its empirical failure is convincing. The maximum that can be said in defense of the EH is that (1) there are short and calm periods in the post-WWII period when at least on the long end of the yield it can't be rejected, or (2) there are different types of departure from the standard neoclassical framework (e.g. time-varying term premia, Peso-problems, habit formation, market microstructure effects), which can produce anomalies very similar to those empirically documented about the behavior of the yield curve (usually meaning the post-WWII U.S. yield curve). The whole story of the yield curve in macroeconomic analysis is basically written in terms of *point estimates* and higher moments of the stochastic processes are completely forgotten (usually melt into the different constants of the models). The behavior of the monetary authority is the center of the vast literature about monetary policy rules, such as Taylor (1993). About the determinants of term premia both theory and empirics are much thinner.

In the finance approach - after the empirical failure of the EH - talking about time-varying term premia (usually interpreted as risk premia) is basically a tautology unless these premia are modeled by putting some structure on their evolution and cross sectional relations. The bulk of the story of the yield curve in finance is basically written in terms of risk and the price of risk. Though the term structure models in finance allow for time varying term premia in the form of ARCH, GARCH, etc. processes, this is only phenomenology: finance never asks about fundamental economic reasons and hence never can answer economic policy questions. In the terminology of Cochrane (2001, p. xiv) finance only can offer models of relative pricing that unlike absolute pricing do not refer to fundamental sources of macroeconomic risk.

As Cochrane says: "The central and unfinished task of absolute asset pricing is to understand and measure the sources of aggregate or macroeconomic risk that drive asset prices. Of course, this is also the central question of macroeconomics, …" . It is this line that we intend to follow.

Our approach to the yield curve is that it is a joint product of the monetary authority and the market and reflects their expectations about the *stochastic processes* governing the economy.

Without modeling financial markets it is not possible to explain higher moments of the stochastic processes driving interest rates, without allowing for effects due to higher moments it is not possible to link properly the short and long end of the yield curve, and without linking properly interest rates with different maturities it is not possible to build a comprehensive monetary macro model.

Dotsey and Otrok (1995) have already forecast: "… a deeper understanding of interest rate behavior will be produced by jointly taking into account the behavior of the monetary authority along with a more detailed understanding of what determines term premia. Reconciling theory with empirical results probably does not require abandonment of the rational expectations paradigm."

As we shall see below, partly we have to contradict: reconciling theory with empirical results probably does require abandonment of the rational expectations paradigm, but not completely. A hypothesis of learning about uncertainty can bring us a long way to explain empirical anomalies. Rational learning is not a completely new idea in the term structure literature. Some authors [e.g. Sack (1998)] investigate the effect of data uncertainty on the optimum monetary policy reaction function, while another line of research [e.g. Sibert (2001)] analyze the effect of unobservable central bank preferences or evaluate specific monetary policy reaction functions by testing the stability of the economy under recursive learning about the reaction function [e.g. Bullard and Mitra (2002)]. In a broader context learning can be interpreted as a special case for time-varying subjective probability adjustment. Mulligan (2004) already studied several aggregate implications of subjective probabilities, such as asset pricing and the relation between consumption growth and capital's return. The novelty of our paper is to propos constant gain learning about the variance of the log pricing kernel, which is in turn determined by the expected variance of the consumption rate of growth and inflation. This type of learning is optimal if conditional variances can be modeled as constants with infrequent but substantial shifts in level.

The main goal of the paper is to show that the long term interest rate does have an economic meaning distinct from both price stability and full employment (though it is not orthogonal to them), namely to measure the market perception of aggregate level of uncertainty or risk in the economy, hence long term interest rates can and have to be an argument in the monetary policy reaction function. If market perception of risk measured as per period conditional variance evolves over time as an AR(1) process, then it can be interpreted either as a constant gain learning or as an optimum weighted average estimate of the conditional variance of the log pricing kernel and hence market expectations not just in terms of conditional mean but also in terms of higher moments of the relevant stochastic processes have to be modeled explicitly.

Section 2 introduces notation and basic definitions in the field of the term structure theory. Section 3 summarizes the main findings of the empirical literature about the U.S. term structure and shows the insufficiency of the assumption of time-varying term premia to explain empirical anomalies. Section 4 introduces a learning hypothesis and section 5 illustrates its use in a consumption-based asset pricing model. Section 6 explores some implications for economic policy and section 7 concludes with possible directions for further research.

2. THE STOCHASTIC PRICING KERNEL AND THE EXPECTATIONS HYPOTHESIS

The expectations hypothesis first formalized by Hicks (1939) is one of the most simple and yet most controversial theory of economic science. In its pure form it states that long term interest rates are equal to the average of the expected future short term interest rates. By the force of no-arbitrage in a deterministic world this is certainly true, but in a stochastic world uncertainty causes systematic distortion in this relation.

In a deterministic world the price of a future payment is determined by the product of the payment and the discount factor:

(1) $$p_t = e^{-nR_{n,t}} X_{t+n}$$

where p_t is the time t price of the deterministic payment X_{t+n} at time t+n and $R_{n,t}$ denotes the time t value of the continuously compounded discount rate for the maturity n.

In a stochastic world the fundamental law of asset pricing states that the price of an asset is the expected value of the product of the payment and the stochastic discount factor:

(2) $$p_t = E_t \{ M_{t,t+n} X_{t+n} \}$$

where $M_{t,t+n}$ is the stochastic discount factor (also called the pricing kernel) between time t and t+n. We have to emphasize that the value of the deterministic discount factor is already known at time t, while the stochastic discount factor at time t is a stochastic variable the actual value of which is known only at time t+n.

In the case of a risk free zero coupon bond the payment is constant across all states of the world $(X_{t+n} = 1)$, hence the time t price of one unit of zero coupon bonds with maturity n is

(3) $$p_t = E_t \{ M_{t,t+n} \}$$

We introduce $m_{t,t+n}$ as the negative logarithm of the stochastic discount factor. Unlike the traditional sign we use the negative logarithm in order to facilitate intuition so that the log pricing kernel moves in the same direction as interest rates:

(4) $$m_{t,t+n} \equiv -\ln(M_{t,t+n})$$

from which follows that the zero coupon interest rate is defined by the equation

(5) $$e^{-nR_{n,t}} = p_t = E_t \{ e^{-m_{t,t+n}} \}$$

The value of the n-period interest rate at time t is thus given by

(6) $$R_{n,t} = -\tfrac{1}{n} \boldsymbol{ln}(p_t) = -\tfrac{1}{n} \boldsymbol{ln}\left(E_t \{ e^{-m_{t,t+n}} \}\right)$$

If the log pricing kernel is normally distributed, then

(7) $$\ln\left(E_t\left\{e^{-m_{t,t+n}}\right\}\right) = E_t\left\{-m_{t,t+n}\right\} + \tfrac{1}{2}Var_t\left\{-m_{t,t+n}\right\} = -E_t\left\{m_{t,t+n}\right\} + \tfrac{1}{2}Var_t\left\{m_{t,t+n}\right\}$$

As it can easily be verified, by the law of one price for any maturity n and k<n

(8) $$m_{t,t+n} = m_{t,t+k} + m_{t+k,t+n}$$

By repeated substitution we can express the n-period log pricing kernel as the sum of future one-period pricing kernels:

(9) $$m_{t,t+n} = \sum_{j=0}^{n-1} m_{t+j+1}$$

where we suppressed the first subscript in the case of one period pricing kernels.

From the definition of the interest rate follows the relation between the interest rates and the normally distributed log pricing kernel:

(10) $$n\left(R_{n,t} + \frac{1}{2}Var_t\left\{\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} m_{t+j+1}\right\}\right) = E_t\left\{m_{t,t+n}\right\}$$

For the sake of simplicity we introduce the notation

(11) $$h_{n,t} = \frac{1}{2}Var_t\left\{\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} m_{t+j+1}\right\}$$

and call it the Jensen term, since it arises in our problem as a result of the Jensen inequality applied for change of ordering between taking expectations and taking logarithm. By the law of iterated expectations

(12) $$E_t\left\{m_{t+k,t+n}\right\} = E_t\left\{E_{t+k}\left\{m_{t+k,t+n}\right\}\right\} = E_t\left\{(n-k)\left(R_{n-k,t+k} + h_{n-k,t+k}\right)\right\}$$

Plugging these into equation (8) and taking time t conditional expectations on both sides we get

(13) $$n\left(R_{n,t} + h_{n,t}\right) = k\left(R_{k,t} + h_{k,t}\right) + (n-k)E_t\left\{R_{n-k,t+k} + h_{n-k,t+k}\right\}$$

This is a risk adjusted form of the expectations hypothesis of the term structure of interest rates (EHTS), which states that long term interest rates are equal to the average of the expected future short term interest rates:

(14) $$R_{n,t} = \tfrac{1}{n}\sum_{j=0}^{n-1} E_t\left\{R_{1,t+j}\right\}$$

implying that for interest rates with different maturities

(15) $$nR_{n,t} = kR_{k,t} + (n-k)E_t\{R_{n-k,t+k}\}$$

The weak form of the expectations hypothesis allows for a constant (time independent) difference between the two sides of the equation. The difference is called the term premium.

If the Jensen terms are constant over time, then the weak form of the expectations hypothesis holds. Moreover, if the Jensen terms are zero, then the strong form of expectations hypothesis holds. If the effect of Jensen terms only could vanish in the complete absence of uncertainty, then there would be no point in testing empirically the strong form. However, this is not the case. The necessary condition for the pure form of the expectations hypothesis to hold is that

(16) $$nh_{n,t} = kh_{k,t} + (n-k)E_t\{h_{n-k,t+k}\}$$

A sufficient condition for this to be true for any pair of maturities n and k is that the Jensen-term is equal across maturities, but follows a martingal process, that is the best possible forecast of any future value is the current value. A special case of this is when the Jensen-term is constant over time, and only an even more special case is when all Jensen-terms are zero, which is the deterministic case.

3. MAIN FINDINGS OF THE EMPIRICAL LITERATURE ON THE US YIELD CURVE

The US yield curve of risk free nominal interest rates is one of the most investigated time series of the world economy. In this section we summarize the main qualitative and quantitative findings, which are relevant from our viewpoint.

3.1. A double affine model

In order to evaluate the stylized empirical findings of the quantitative literature in a unified framework, first we present a general affine model of the yield curve, where both the short term (one-period) interest rate and the Jensen-terms (expected conditional variance of the log pricing kernel for any horizon) are affine functions of the $\mathbf{z}_t$ vector of state variables, which in turn follow a zero mean VAR(1) process:

(17) $$\begin{aligned} R_{1,t} &= g(1) + \gamma'(1)\mathbf{z}_t \\ h_{k,t} &= d(k) + \delta'(k)\mathbf{z}_t \\ \mathbf{z}_t &= \mathbf{A}\mathbf{z}_t + \mathbf{u}_{t+1} \\ \mathbf{u}_{t+1} &\sim N(\mathbf{0}, \mathbf{\Sigma}) \end{aligned}$$

where $\mathbf{A}$ is the invertible matrix of autoregressive coefficients, at most one of its eigenvalues is on the complex unit circle, all the others are inside, $\mathbf{u}$ is the vector of innovations and $\mathbf{\Sigma}$ is the covariance matrix of the innovations with full rank (that is all the state variables are truly stochastic variables, not constant). State variables can be e.g. consumption growth and inflation in a demeaned form. The short rate (one-period interest rate) is a state variable if

$\gamma'(1)$ is a selection vector. (One element 1, all others zero.) Without loss of generality we choose the first selection vector $\gamma(1)=\mathbf{e}_1$, hence

(18) $R_{1,t} = \overline{R_1} + e_1' z_t$

where $\overline{R_1}$ is the unconditional mean of the one-period interest rate (if it exists).

From this follows that both long-term interest rates and log pricing kernels are also affine functions of the state variables.

(19)
$$R_{n,t} = g(n) + \gamma'(n)\mathbf{z}_t$$
where
$$\gamma'(n) = -\delta'(k) + [\mathbf{e}_1' + \delta'(1)]\tfrac{1}{n}\sum_{k=0}^{n-1}\mathbf{A}^k = -\delta'(k) + [\mathbf{e}_1' + \delta'(1)]\tfrac{1}{n}(\mathbf{I}-\mathbf{A})^{-1}(\mathbf{I}-\mathbf{A}^{-n})$$
$$g(n) = \overline{R}_1 + d(1) - d(n) = \overline{R}_n$$

where the closed form of the summations holds if the vector process of state variables is stationary.

This form makes explicit the interaction of monetary policy and market expectations. Namely monetary policy can be represented in most industrial countries by the short-term interest rate since it is its most important instrument to influence economic activity and the rate of inflation. The shortest-term interest rate is the one-period rate determined by $\gamma(1)$ and $g(1)$. The law of motion of the short-term interest rate can be interpreted as the monetary policy reaction function and the first element of $\mathbf{u}_{t+1}$ can be interpreted as a monetary policy shock:

(20) $R_{1,t+1} = \overline{R_1} + \mathbf{e}_1'\mathbf{A}z_t + \mathbf{e}_1'\mathbf{u}_{t+1}$

The coefficients $d(n)$ and $\delta(n)$ determine the Jensen-terms which measure market expectations about uncertainty over the n-period horizon. Long-term interest rates are determined by market expectations on conditional variance of the log pricing kernel that is the expected volatility of macroeconomic variables such as the growth rate of consumption and the rate of inflation.

One could argue that there is no point in separating the Jensen terms from the interest rates. After all, if both are affine function of the state variables, then their sum, the log pricing kernel will also follow an affine model:

As a combination of the short-term interest rate and the one-period Jensen-term we get the expression for the expected value of the log pricing kernel, which is again an affine function of the state variables:

(21)
$$E_t\{m_{t+1}\} = f + \phi'\mathbf{z}_t$$
where
$$\phi' = \mathbf{e}_1' + \delta'(1)$$
$$f = \bar{R}_1 + d(1)$$

From this follows that the error term η_{t+1} in the equation

(22) $$m_{t+1} = E_t\{m_{t+1}\} + \eta_{t+1}$$

is a martingal difference sequence. The log pricing kernel m_t can also be expressed as an affine function of the vector of state variables $\mathbf{z}_t$

(23)
$$m_t = f + \phi'\mathbf{A}^{-1}\mathbf{z}_t$$
and
$$\eta_{t+1} = \phi'\mathbf{A}^{-1}\mathbf{u}_t$$

The main advantage of the separation is on one hand that interest rates unlike pricing kernels are more or less directly observable and on the other hand the use of the below mentioned empirical results, namely that term premia can be modeled as univariate functions of a highly persistent latent variable and this latent variable overwhelmingly determines the behavior of long term interest rates. To illustrate the functioning of the model we present it without macro variables, so there is only the short term interest rate and the Jensen factor will be specified, the other elements of $\mathbf{z}_t$ will remain general variables, observable or latent.

An important weakness of our double affine model is that it doesn't allow for the zero lower bound on nominal interest rates. Since due to real world economic policy problems this lower bound became extremely important in recent years, this is certainly an issue, which has to be addressed in future research.

3.2. Unit root and cointegration

US nominal interest rates are either integrated or at least near integrated processes, which are at least very close to be pair wise cointegrated with the cointegrating vector (1, -1). Gil-Alana (2004) supports the hypothesis of a unit root in US long-term interest rates in the period 1940-2000. Engsted and Tangaard (1994) come to the clear conclusion of pair wise cointegration with the cointegrating vector (1,-1). Österholm (2004) tests the relevance of the unit root hypothesis for US unemployment, real exchange rate, nominal interest rate and inflation by defining them as a panel and thereafter applying three different panel unit root tests. Based on monthly data from 1960 to 2002 his results indicate that all four series are generated by highly persistent, but stationary, processes.

If interest rates are pair wise cointegrated with the cointegrated vector (1,-1), then by definition risk premia must be stationary. The most important problem in deciding whether interest rates are stationary or not is the structural break at the beginning of the Volcker era (end of 1970s-early 1980s). Bernanke and Blinder (1992) has put forward the since than widely accepted idea that the best measure of the stance of US monetary policy is the Federal Funds rate, the O/N rate at which commercial banks trade federal reserve funds among themselves. This has been true for at least the last two decades and probably even earlier, except for the above mentioned period, when the Fed adopted a different monetary policy operating procedure targeting the quantity of free reserves of commercial banks instead of the federal funds rate.

A few articles directly test the role of the holding period excess return on long term bonds. The holding period excess return on a zero coupon bond with maturity *n* is by definition:

(24) $$X_{n,t} \equiv nR_{n,t} - (n-1)R_{n-1,t+1} - R_{1,t}$$

Substituting the definition from the closed form solution gives

(25) $$E_t\{X_{n,t}\} = [n\delta'(n) - (n-1)\delta'(n-1)\mathbf{A} - \delta'(1)] - [nd(n) - (n-1)d(n-1) - d(1)]$$

Evans and Lewis (1994) test whether stationary risk premia can alone explain the behavior of excess returns to long bonds relative to rolling over short rates. They reject this hypothesis using U.S. T-bill returns. They then show that either permanent shocks to the risk premia and/or rationally anticipated shifts in the interest rate process could produce anomalous results. Tzavalis and Wickens (1995) examine the persistence of the volatility of the risk premia for excess holding period returns using a GARCH-M model of the conditional variance. They find that the high degree of persistence cannot be sustained once allowance is made for a structural break in the unconditional variance caused by a change in the operation of US monetary policy during 1979-1982.

In the frame of our VAR model we can summarize these results as follows:

Among the factors there is one, which produces the (near) unit root result in all the interest rates and its loading must be the same for all maturities in order to have the cointegrating vector (1,-1). Either this factor also enters the term premium, or there must be a second factor which is close to unit root as well.

3.3. Stylized empirical facts about the Jensen-term

As Tzavalis (2003) shows, a univariate AR(1) representation of the excess return is a satisfactory model of postwar US data.

Bams and Wolff (2003) estimate excess returns in a panel data approach. They assume that the excess holding period returns for yields with different time-to-maturity are driven by the same dynamics and impose a time-invariant one-factor model in time dimension. They also incorporate the maturity dimension of risk premia. The excess holding period returns of the long-term yield serves as a base case. The risk premia of all other yields are assumed to be related to this process through a scaling factor. Their model is rather specific, but the results are sharp enough. (Here we use our notation introduced above.)

$$
\begin{aligned}
& nR_{n,t} = \beta\left[(n+1)R_{n+1,t} - R_{1,t}\right] - Z(n)\xi_{t+1} + \eta_{n,t+1} \\
& \text{where} \\
& \xi_{t+1} = (1-\rho)\mu + \rho\xi_t + \nu_{t+1} \\
\textbf{(26)}\quad & \boldsymbol{\eta}_{t+1} \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \\
& \nu_{t+1} \sim N\left(0, \sigma_\nu^2\right) \\
& (\boldsymbol{\eta}_{t+1}, \nu_{t+1}) \text{ independent} \\
& \boldsymbol{\Sigma}(i,j) = \omega^2 \frac{\phi^{\left|n_i - n_j\right|}}{\left(n_i n_j\right)^d}
\end{aligned}
$$

where ξ_{t+1} is the time varying term premium, which follows an AR(1) process with autocorrelation coefficient ρ and innovation variance σ^2, while the error terms in the expectations hypothesis equations for different maturities $\eta_{n,t+1}$ follow a multivariate normal distribution with zero mean and $\boldsymbol{\Sigma}$ covariance matrix. The last line specifies covariance between the error terms for different maturities, where ω, ϕ and d are scalar parameters. Substituting Jensen-terms for the interest rates and rearranging terms we get

$$
\textbf{(27)}\quad -Z(n)\xi_{t+1} + \eta_{n,t+1} = (1-\beta)nR_{n,t+1} - \beta\left[(n+1)h_{n+1,t} - nh_{n,t+1} - h_{1,t} - \varepsilon_{n,t+1}\right]
$$

From our point of view the key result is that $Z(n)$ is a strictly increasing, asymptotically linear function of the maturity n. This is true even if the slope coefficient β is forced to be equal to 1 (as it is required by the expectations hypothesis). Plugging in for the Jensen-terms from (17) and taking expectations on both sides we get

$$
\textbf{(28)}\quad
\begin{aligned}
& -Z(n)\left[\rho\xi_t + (1-\rho)\mu\right] = -\left[(n+1)h_{n+1,t} - nh_{n,t+1} - h_{1,t}\right] = \\
& = -\left[(n+1)\boldsymbol{\delta}'(n+1) - n\boldsymbol{\delta}'(n)\mathbf{A} - \boldsymbol{\delta}'(1)\right]\mathbf{z}_t - \left[(n+1)d(n+1) - nd(n) - d(1)\right]
\end{aligned}
$$

Equating terms and accepting the result of Tzavalis (2003) concerning the univariate representation of the excess return implies that

$$
\textbf{(29)}\quad Z(n)\rho = (n+1)D(n+1) - nD(n)\mathbf{A}_{jj} - D(1)
$$

where ξ_t is the jth element of $\mathbf{z}_t$ and hence $A_{jj} = \rho$. If we accept the result of Bams and Wolff (2003) that $Z(n)$ is an affine function of n in the form[2]

$$
\textbf{(30)}\quad Z(n) = an + b
$$

---

[2] The theoretical hypothesis of the holding period excess return as an affine function of a single, highly persistent AR(1) process has already been used in Tzavalis and Wickens (1997) and Romhányi (2002) in order to explain the empirical failure of the EH.

then the functional form of the multi-period Jensen-term has to be

$$\textbf{(31)} \quad D(n) = \frac{\rho a}{1-\rho} - \frac{b}{n}$$

If $b = 0$, as it is the case in Bams and Wolff (2003), then $D(n)$ is constant that is independent of the maturity and the Jensen-term asymptotically (as n goes to infinity) takes the form:

$$\textbf{(32)} \quad h_{n,t} \overset{n\to\infty}{\cong} d_n + \frac{\rho a}{1-\rho}\xi_t$$

In words, the long term Jensen-term is the sum of two parts, one of which depends only on maturity[3] and the other only on time. The factor determining the Jensen-terms will be called the Jensen-factor and let us put it on the second place in the vector of state variables. From this follows that

$$\textbf{(33)} \quad h_{n,t} \overset{n\to\infty}{\cong} d_n + \frac{\rho a}{1-\rho}\mathbf{e}_2\mathbf{z}_t$$

A further corroboration of the eminent role of the Jensen-factor in shaping the yield curve can be found in Cochrane and Piazzesi (2005). They investigate the time variation in expected excess bond returns. They run regressions of annual excess returns on forward rates and find that a single factor predicts 1-year excess return on 1-5 year maturity bonds with an R^2 up to 43%. This single factor is a tent-shaped linear function of forward rates. The return-forecasting factor has a clear business cycle correlation: expected returns are high in bad times, and low in good times, and the return-forecasting factor also forecasts stock returns, suggesting a common time-varying premium for real interest rate risk. The return-forecasting factor is poorly related to level, slope and curvature movements in bond yields. Therefore, it represents a source of yield curve movements not captured by most term structure models. Though the return-forecasting factor accounts for more than 99% of the time-variation in expected excess bond returns, they find additional, very small factors that forecast equally small differences between long term bond returns, and hence statistically reject a one-factor model for expected returns.

Though the two models in Bams and Wolff (2003) (with and without the restriction $\beta = 1$) give the same asymptotic result, in the case, where the interest rate enters the equation, the statistical properties of the regression are much clearer and better. If beyond the main latent factor the long term interest rate also has an effect on the holding period excess return, then how can the excess return be one-dimensional?

---

[3] Since the equation of Bams and Wolff (2003) doesn't contain a both time and maturity independent constant we can't exactly determine the functional form of $d(n)$.

3.4. The behavior of long term interest rates

As we saw in the general inference of interest rates from the log pricing kernel, longer term interest rates depend on the conditional mean and conditional variance of the one period log pricing kernel over the period to maturity of the interest rate. Under the rather weak condition of ergodicity the conditional mean and conditional variance of the infinite sum converges to the unconditional mean and variance. If there are no structural breaks in the stochastic processes governing the economy, very long term interest rates should be constant, or have at least much lower volatility, than short term interest rates.

Compared to this theoretical requirement, the empirical literature reports about excess volatility in long term interest rates. [e.g. Shiller and McCulloch (1990)]

Dybvig, Ingersoll and Ross (1996) states that long term zero-coupon interest rates, if they exist at all, can never fall. Though their proof of the theorem had to be corrected by McCulloch (2000), their basic argument is that if there is a lower bound, deterministic or stochastic of long term interest rates, then rational investors can construct an arbitrage portfolio. The reasoning hinges on three assumptions: rationality, full information and no transactions costs. McCulloch (2000) shows that the introduction of transactions costs can invalidate the result, because any positive constant bid-asked spread results in infinitely diverging bid and asked zero-coupon yield curves rendering the average zero coupon yield curve indeterminate. Nevertheless, based on bid-asked mean price quotations that do not take transactions costs into account McCulloch and Kochin (2000) find no evidence that the estimated forward rate beyond 30 years is nondecreasing over time, or even has lessened variance.

Jordà and Salyer (2003) show that greater uncertainty about monetary policy can lead to a decline in nominal interest rates.

Ang and Piazzesi (2003) describe the joint dynamics of bond yields and macroeconomic variables in a Vector Autoregression, where identifying restrictions are based on the absence of arbitrage. Using a term structure model with inflation and economic growth factors, together with latent variables, they investigate how macro variables affect bond prices and the dynamics of the yield curve. They find that the forecasting performance of a VAR improves when no-arbitrage restrictions are imposed and that models with macro factors forecast better than models only with unobservable factors. Variance decompositions show that macro factors explain up to 85% of the variation in bond yields. Macro factors primarily explain movements at the short end and middle of the yield curve while unobservable factors still account for most of the movement at the long end of the yield curve.

Based on the above inference we can recognize at least one of the unobservable factors determining the long end of the yield curve in the Jensen-factor.

According to the closed form solution for the very long term interest rates in the general VAR-model we have for the time varying part

$$\textbf{(34)} \quad \lim_{n\to\infty} \gamma'(n)\mathbf{z}_t = \lim_{n\to\infty}\left[-\delta'(k)+\left[\mathbf{e}_1'+\delta'(1)\right]\tfrac{1}{n}\sum_{k=0}^{n-1}\mathbf{A}^k\right]\mathbf{z}_t = \left[-\frac{a\rho}{1-\rho}\mathbf{e}_2' + \left(\mathbf{e}_1' + \frac{a\rho}{1-\rho}\mathbf{e}_2'\right)\mathbf{A}^\infty\right]\mathbf{z}_t$$

By the assumption that at most one of the eigenvalues of **A** is on the complex unit circle, at most one diagonal element of $\mathbf{A}^{\infty}$ can be 1, all the other elements must be zero. There are thus four main cases for the value of $\mathbf{A}^{\infty}$, which we can characterize by the vector of diagonal elements.

	$\mathbf{A}^{\infty}$	$\lim_{n\to\infty} \gamma'(n)$	$\lim_{n\to\infty} \gamma'(n)\mathbf{z}_t$
1.	$\langle 0 \rangle$	$-\frac{a\rho}{1-\rho}\mathbf{e}_2$	$-\frac{a\rho}{1-\rho}\xi_t$
2.	$\langle \mathbf{e}_1 \rangle$	$\mathbf{e}_1 - \frac{a\rho}{1-\rho}\mathbf{e}_2$	$R_{1,t} - \frac{a\rho}{1-\rho}\xi_t$
3.	$\langle \mathbf{e}_2 \rangle$	0	0
4.	$\langle \mathbf{e}_{S>2} \rangle$	$-\frac{a\rho}{1-\rho}\mathbf{e}_2$	$-\frac{a\rho}{1-\rho}\xi_t$

Table 1 The main types of the VAR coefficient matrix and the long end of the yield curve

In the 1st case there is no exact unit root in the process (the empirical failure to reject the hypothesis is either due to structural breaks or to an eigenvalue very close to unity) and long term interest rates are only determined by the Jensen-factor.

In the 2nd case the short term interest rate imposes the unit root on all the other interest rates, including the very long ones. Since this only means a parallel level shift in the whole yield curve, the term premium is still determined by the univariate process of the Jensen-factor. Empirically this case is rather problematic, because any change in the short term interest rate should always cause a parallel shift in the whole yield curve, which doesn't seem to be the case in reality. (See e.g. Campbell (1995)) Moreover this contradicts the empirical result of Ang and Piazzesi (2003) that unobservable factors account for most of the movement at the long end of the yield curve.

In the 3rd case, when the Jensen-factor follows a unit root process, long term interest rates are constant. This case can easily be excluded, because a unit root in the Jensen-factor is incompatible with any finite value for the Jensen-terms due to the zero denominator.

In the 4th case, when some other factor (S>2) imposes the unit root on the yield curve, very long term interest rates are still determined by the Jensen-factor. This is in contradiction with the at least approximate empirical result about a (1,-1) cointegrating vector for any pair of interest rates.

All together we are left with the first case when there is no exact unit root in the whole vector process but the Jensen-factor is highly persistent rendering both the interest rates and the term premia highly persistent. (Automatically this also means that the log pricing kernel is stationary as well.) This is in line with the empirical result of Österholm (2004), and contradicts the result of Gil-Alana (2004). Moreover, since long term interest rates are driven by the Jensen-factor, they preserve the one-dimensional property of the holding period excess return when entering the regression in Bams and Wolff (2003).

3.5. Tests of the expectations hypothesis

Some papers investigate the EH in different sub samples. Favero and Mosca (2000) investigates the reaction of the short end of the US yield curve and finds that the expectations hypothesis cannot be rejected in periods of low uncertainty on monetary policy. Christiansen (2003) tests the expectations hypothesis using forward-rate regression and concludes that for short data sets the EH is close to being accepted, but for long data sets it is rejected. Lanne (1999a) also finds that high persistence in yield spreads can cause the statistical rejection of the EH, but the degree of persistence in US interest rates in the period 1952-1991 has varied significantly. Tzavalis and Wickens (1997) show that the empirical failure of the EH might also be explained by a time-varying term premium (expected holding period excess return) that is correlated with the term spread (the difference between long and short term interest rates).

In our approach equation (16) ensures that over short periods of relative stability the expectations hypothesis must be a good approximation at least at the long end of the yield curve.

In Campbell and Shiller (1991) there are two regressions to test the expectations hypothesis of the term structure of US nominal interest rates for the period 1952-1987. The second one is an approximation of the first one for the long and of the yield curve, where *n* is large. Since the most striking result there is the slope coefficients along a strait line with negative slope (instead of being constant 1), we are interested only in the long end of the yield curve, so calculate the theoretical value of the slope coefficient in the second regression:

(35) $$R_{n,t+k} - R_{n,t} = \alpha(n,k) + \beta(n,k)\frac{k}{n-k}\left(R_{n,t} - R_{k,t}\right) + \varepsilon_{t+k}$$

The weak form of the EH implies that $\beta(n,k) = 1$ for any $0 < k < n$.

The stylized empirical result of Campbell and Shiller (1991) is that the slope coefficient is an asymptotically affine function of the maturity *n* (though the limit depends on the sample period):

(36) $$0 < \left|\lim_{n\to\infty}\frac{\partial}{\partial n}\hat{\beta}(n,k)\right| < \infty$$

As it is shown in Appendix A this implies that

$$0 < \lim_{n\to\infty} \| \gamma(n) \| < \infty$$

(37) and

$$0 < \lim_{n\to\infty} \| \delta(n) \| < \infty$$

which means that both time varying long term interest rates and time varying long term Jensen-terms are needed for the stylized empirical results of Campbell and Shiller (1991) to show up. This is in clear contradiction with the traditional analysis of the yield curve, but based on the above analysis about the behavior of long term interest rates this is exactly what we could have expected.

4. A RATIONAL LEARNING HYPOTHESIS OF THE TERM STRUCTURE

The crucial result from contrasting theory with empirical findings is the clear contradiction between the behavior of long term interest rates implied by the convergence to a constant in the pricing kernel model and the permanent change implied by the Campbell-Shiller regressions. If the n-period Jensen-term $h_{n,t}$ is the average variance of expected one period log pricing kernels according to equation (11), then (after having established above the stationarity of the log pricing kernel) under the rather weak condition of ergodicity, the infinite maturity Jensen-term must converge to the constant value of the unconditional variance and hence it can't change over time. On the contrary, in order to produce asymptotic behavior of the Campbell-Shiller regression coefficients, even very long term interest rates must have some variance. In other words, stochastics is certainly needed to explain the failure of the pure form of the EH. Moreover, time varying term premia are certainly needed to resolve the failure of the weak form of the EH, but they are still not enough to produce the Campbell-Shiller results. In a mathematical form the stochastic process of the log pricing kernel should fulfill the following, impossible requirement at least for medium and long maturities and especially for infinity:

$$\frac{1}{2}\mathrm{Var}_t\left\{\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} m_{t+j+1}\right\} = h_{n,t} \overset{n\to\infty}{\cong} d(n) + \frac{\rho a}{1-\rho}\xi_t$$

(38) $$\xi_t = \rho\xi_{t-1} + \nu_t$$

$$\nu_t \sim N\left(0, \sigma_\nu^2\right)$$

Assuming a time-varying price of risk can't solve the problem, because the risk premium is the product of the risk and the price of risk. If both processes (the level of risk and its price) are stationary, then the infinite maturity risk premium still has to be constant over time.

Our proposal to resolve the problem is the introduction of rational learning instead of rational expectations on behalf of financial market participants. By this assumption the two different concepts of the Jensen-term become distinct. *Ex ante* Jensen-terms are defined as the conditional variance of the log pricing kernel, while *ex post* Jensen terms are derived from the term premia, which are econometrically estimated from past data.

The simplest way we can allow for learning is to conjecture that the theoretically constant value of the conditional variance of some specific type of shocks (e.g. monetary policy, productivity, etc.) is unknown. It has to be seen as a simplification of a broader set of models, where productivity shocks are heteroskedastic, (possibly ARCH or some Markov-process with regime-shifts), but their variance is stationary. Moreover it is an utmost simplification to analyze a model like our, where only the value of one single parameter is unknown, but all the others are perfectly measured common knowledge.

If the volatility of the shocks is unknown, then it must be estimated from empirical data. Moreover, there might be changes in this constant over sufficiently long periods, but modeling these changes is even less feasible, so the best way is to use a rolling window estimate where recent data have higher weight but more distant data also enter the calculation though with smaller weight. Already Muth (1960) showed that exponential weighting is optimal if the time series the current value of which has to be forecast is the composite of a random walk and an iid random variable. The weights decrease the slower the smaller is the variance of the permanent shocks compared to the variance of the transitory shocks. Foster and Nelson (1996, theorem 5) also found in a continuous time setting that the asymptotic variance-minimizing backward looking weight function (the only possible way to estimate current value) to estimate conditional variance is exponential. Exponential weighting of past instantaneous values is equivalent with an AR(1) process, since

(39) $$\hat{\sigma}^2_{\varepsilon,t} = \rho\hat{\sigma}^2_{\varepsilon,t-1} + (1-\rho)\varepsilon^2_t$$

with $0 \le \rho < 1$ is equivalent with

(40) $$\hat{\sigma}^2_{\varepsilon,t} = (1-\rho)\sum_{j=0}^{\infty}\rho^j\varepsilon^2_{t-j}$$

In the literature of rational learning this could also be called a constant gain learning rule (see. e.g. Evans and Honkapohja, 2001, p.48).

At first site it looks as if the shocks followed an ARCH-process, but this is not the case. The shocks are homoskedastic, only the market estimate of the constant variance evolves over time as an AR(1) process and moreover this is only true from the outside researcher's point of view. If one asked a market participant what he or she thinks about the variance of the shocks, the answer would be that the shocks are homoskedastic, with the variance $\hat{\sigma}^2_{\varepsilon,t}$. If the researcher asked the market participant about his or her next day's estimate, the answer would be: "It will almost certainly change, but I don't know, in what direction, so my best estimate is still $E_t\left\{\hat{\sigma}^2_{\varepsilon,t+1}\right\} = \hat{\sigma}^2_{\varepsilon,t}$". So from the investor's point of view shocks are permanent since they alter his estimate of expected variance at any horizon. Put in another way, the investor looking forward into the future will take decisions as if the weak form of the expectations hypothesis was true, and still, any econometric test based on long enough time series of historical data will reject the hypothesis.

If the shock with unknown variance is symmetrically distributed, then its current value and its square (the measure of the "instantaneous" variance) will be uncorrelated. Since current "instantaneous" values of shocks are extremely difficult to measure the process of learning will leave ample room for non-linearities or even Peso-problems. This is something already well documented in the literature. Ang and Bekaert (2000), Ang and Bekaert (2001) and Bansal and Zhou (2002) attribute the empirical failure of the expectations hypothesis to regime shifts in interest rates, while several papers, such as Lanne (1999b), Bekaert, Hodrick and Marshall (2001) show that even Peso-problems might be enough to produce some of the documented anomalies. Empirically this will add to the degree of exogeneity of the stochastic process with respect to fundamental macroeconomic variables, such as inflation, consumption or short term interest rates. This exogeneity of a new factor in connection with volatility beyond the usual three factors allowed for in traditional affine term structure models has already been raised by Collin-Dufresne, Goldstein and Jones (2004).

Probably everybody agrees that for instance the October 1987 crash in the American stock market is still alive in the memory of most of today's stock traders. If this is true, then the weight put on more distant experiences can't decay too fast and hence the time-series of the Jensen-factor must be extremely persistent.[4] This fact might account for the empirical results of Evens and Lewis (1994), who found that stationary term premia can't explain the behavior of excess returns in the bond market.

5. Illustrative implementation in a consumption-based asset pricing model

Starting from the idea of rational learning we specify in more detail a possible implementation of our hypothesis. The starting point is the exchange economy analyzed by Lucas (1978) but amended with production and an infinite array of risk free bonds. It is only meant to be an illustrative example and certainly not a definitive model of the whole behavior of the yield curve.

To solve the problem of the expectations hypothesis and excessive variability of the risk premia, a line of research explores the possible effects of alternative utility functions, such as external habit formation (Campbell and Cochrane, 1999) and non-expected utility [Gregory and Voss (1991) for both effects for the Canadian term structure]. The general conclusion is that neither preference structure is able to mimic satisfactorily the magnitude or the variability of the risk premia, so there is no strong reason to depart from the most traditional neo-classical model of human behavior. Regarding the overwhelming evidence about human behavior not in line with the standard neo-classical assumptions (fully rational maximization of exponentially discounted expected utility over current and future consumption) we have to refer to the argument put forward by Merton and Bodie (2004):"When particular transaction costs or behavioral patterns produce large departures from the predictions of the ideal 'frictionless' neoclassical equilibrium for a given institutional structure, new institutions tend to develop that partially offset the resulting inefficiencies. In the longer run, after institutional structures have had time to fully develop, the predictions of the neoclassical model will be

---

[4] E.g. to have a 5% relative weight on data from 20 years ago compared to current data, first order autocorrelation of monthly data must be above 0.987. In order to reject a unit root at 5% level, when the true autocorrelation is 0.987, one needs a sample size of more 640, which is more than 50 years.

approximately valid for asset prices and resource allocations." If this argument is valid anywhere at all, then the US government bond market is among the first candidates.

5.1. Consumption based government bond pricing

We model investors by a time separable, exponentially discounted expected utility function defined over current and future values of consumption,

$$\textbf{(41)} \quad U(C_t, C_{t+1}, C_{t+2}, \ldots) = E_t \sum_{j=0}^{\infty} \beta^j u(C_{t+j})$$

where C_t is consumption at time t, β is the subjective discount factor, $u(\cdot)$ is the strictly increasing, twice continuously differentiable and concave instantaneous utility function and $E_t\{\cdot\}$ denotes mathematical expectations conditional on information available at time t.

We assume that the investor can freely buy or sell as much of risk-free zero coupon bonds of any maturity, as he or she wishes. The time t nominal price of a risk-free bond maturing at time $t+k$ is $p_{k,t}$. There is a single aggregate commodity in the economy used for both consumption and production, the price of which is denoted by P_t. In any time period t there are three purposes the investor can use his money for: consumption (C_t), real fixed investment (I_t) and zero coupon bonds ($q_{k,t}$). He has money from two sources: current income Y_t and maturing bonds. (Selling a bond before maturity is by the law of one price equivalent with a negative investment into a bond with the maturity equal to the remaining time to maturity of the longer term bond.)

Real fixed investment is the only input of production of the single composite good via a stochastic (but analytically well behaving) production function $f(\cdot)$. Capital (K_t) depreciates at the constant rate δ and the logarithm of productivity shocks are independently and normally distributed with zero mean. The domestic investor's problem can thus be summarized in the formulae:

$$E_t \sum_{j=0}^{\infty} \beta^j u(C_{t+j}) \rightarrow \boldsymbol{max}$$

$$\text{s.t.} \quad \forall j \geq 0$$

$$\textbf{(42)} \quad P_{t+j}C_{t+j} + P_{t+j}I_{t+j} + \sum_{j=0}^{\infty} p_{k,t+j}q_{k,t+j} = P_{t+j}Y_{t+j} + \sum_{j=0}^{\infty} q_{k,t+j-k}$$

$$Y_{t+j} = f(K_{t+j-1}) * \boldsymbol{exp}(\varepsilon_{t+j})$$

$$\varepsilon_{t+j} \sim \text{NID}(0, \sigma_\varepsilon^2)$$

$$K_{t+j} = (1-\delta)K_{t+j-1} + I_{t+j}$$

As it is inferred in the Appendix B, the yield curve is given by the equations

$$\textbf{(43)} \quad R_{n,t} = -\frac{1}{n}\ln p_{n,t} = -\frac{1}{n}\ln E_t\left\{\beta^n \frac{u'(C_{t+n})}{u'(C_t)}\frac{P_t}{P_{t+n}}\right\}$$

$$\textbf{(44)} \quad \beta^n \frac{u'(C_{t+1})}{u'(C_t)} = \frac{1}{1-\delta+f'(K_t)*\exp(\varepsilon_{t+1})}$$

where $p_{n,t}$ denotes the time t price of one unit of risk free bond maturing at time t+n and $R_{n,t}$ is its yield to maturity. The expression within parenthesis in the first equation is the stochastic pricing kernel between time t and t+n, the logarithm of which will be denoted again by $m_{t,t+n}$.

5.2. Linearization

For the sake of analytical tractability we make a number of assumptions: power utility function, AK technology, multivariate lognormal distribution and conditional homoskedasticity.

Power utility function[5] implies that the log pricing kernel is an affine function of consumption rate of growth and inflation

$$\textbf{(45)} \quad m_{t,t+1} = -\ln\left[\beta^n \frac{u'(C_{t+1})}{u'(C_t)}\frac{P_t}{P_{t+1}}\right] = -\ln\beta + \gamma\Delta c_{t+1} + \pi_{t+1}$$

where $\Delta c_{t+1} \equiv \ln C_{t+1} - \ln C_t$ is the growth rate of consumption and π_{t+1} is the rate of inflation between time t and t+1. Compared to the traditional notation here again we changed the sign of the log pricing kernel in order to strengthen intuition, since the "original" pricing kernel is related to the one period bond price and moves in the opposite direction as the short term interest rate. For further use we keep the notation of the one period log pricing kernel $m_{t+1} \equiv m_{t,t+1}$.

AK technology implies that the equilibrium rate of consumption growth does not depend on the stock of capital:

$$\textbf{(46)} \quad -\ln\beta + \gamma\Delta c_{t+1} = \ln(1-\delta + A*\exp\varepsilon_{t+1})$$

and by a linear approximation of the RHS of the equation[6]

---

[5] The power utility function with the parameter of relative risk aversion γ linearizes the logarithm of the marginal utility of consumption: $u(C_t) = \frac{(C_t)^{1-\gamma}-1}{1-\gamma} \quad \Rightarrow \quad \ln u'(C_t) = -\gamma \ln C_t \equiv -\gamma\, c_t$

(47) $c_{t+1} = c_t + \frac{1}{\gamma}\ln[\beta(1-\delta+A)] + \frac{1}{\gamma}\frac{A}{1-\delta+A}\varepsilon_{t+1}$

which means that the expected rate of consumption growth is constant over time or the logarithm of the level of consumption can be modeled as a random walk with a deterministic drift. This equation is in close connection with the result of Hall (1978), where the possible "random walk" property of consumption has first been derived.

If we use a quadratic approximation[7], then the expected rate of consumption growth instead of being constant depends on the conditional variance of the productivity shock:

(48) $E_t\{c_{t+1}\} = c_t + \frac{1}{\gamma}\boldsymbol{ln}[\beta(1-\delta+A)] + \frac{1}{\gamma}\frac{1}{2}\frac{(1-\delta)A}{(1-\delta+A)^2}\sigma_\varepsilon^2$

Lognormal distribution implies that the difference between the logarithm of the expected value and the expected value of the logarithm is half of the variance of the logarithm:

(49) $R_{n,t} = E_t\left\{\frac{1}{n}\sum_{j=1}^{n} m_{t+j}\right\} - \frac{1}{2}Var_t\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n} m_{t+j}\right\}$

From equation (45) we have for the conditional expectations term

(50) $E_t\left\{\frac{1}{n}\sum_{j=1}^{n} m_{t+j}\right\} = -\ln\beta + \gamma E_t\left\{\frac{1}{n}\sum_{j=1}^{n} c_{t+j}\right\} + E_t\left\{\frac{1}{n}\sum_{j=1}^{n} \pi_{t+j}\right\}$

For the conditional variance term we need the assumption of conditional homoskedasticity of the productivity shocks:

(51) $\mathrm{Var}_{t+j}\{c_{t+j+1}\} = \left(\frac{A}{1-\delta+A}\right)^2 \mathrm{Var}_{t+j}\{\varepsilon_{t+j+1}\} = \left(\frac{A}{1-\delta+A}\right)^2 \sigma_\varepsilon^2 \quad \forall j \geq 0$

In order to close the model we need a law of motion for the rate of inflation. This can be specified in several ways such as motivated by some monetary policy rule (quantity of money or interest rate rule) that implies the required law of motion or it can be based on some aggregate supply curve relating output and price changes or it can be some autonomous

---

6 To linearize the effect of productivity shocks we use the quadratic approximation of ε_{t+1} around its mean:

$$\ln(1-\delta+A*\exp\varepsilon_{t+1}) = \ln(1-\delta+A) + \frac{A}{1-\delta+A}\varepsilon_{t+1}$$

7 $\ln(1-\delta+A*\exp\varepsilon_{t+1}) = \ln(1-\delta+A) + \frac{A}{1-\delta+A}\varepsilon_{t+1} + \frac{1}{2}\frac{(1-\delta)A}{(1-\delta+A)^2}\varepsilon_t^2$

stochastic process based on price setting behavior of firms as a function of previous and expected price changes, etc. For the sake of simplicity we don't specify the model further, just assume that the one-period ahead forecast error of the rate of inflation is also homoskedastic:

$$\textbf{(52)} \quad \mathrm{Var}_{t+j}\{\pi_{t+j+1}\} = \mathrm{Var}_{t+j}\{\zeta_{t+j+1}\} = \sigma_\zeta^2 \quad \forall j \geq 0$$

where ζ_t can be interpreted as nominal shocks as opposed to ε_t, which have the nature of a real shock. If nominal and real shocks are independent[8], then the n-period conditional variance term can be linearly approximated as an affine function of the two one-period conditional variances[9]:

$$\textbf{(53)} \quad Var_t\left\{\frac{1}{n}\sum_{j=1}^{n} m_{t+j}\right\} = s(n) + s_\varepsilon(n)\gamma^2\sigma_\varepsilon^2 + s_\zeta(n)\sigma_\zeta^2$$

To sum up, our closed form solution for the n-period zero coupon interest rate is:

$$\textbf{(54)} \quad \mathrm{R}_{n,t} = -\boldsymbol{ln}\beta + \gamma \mathrm{E}_t\left\{\frac{1}{n}\sum_{j=1}^{n}\Delta c_{t+j}\right\} + \mathrm{E}_t\left\{\frac{1}{n}\sum_{j=1}^{n}\pi_{t+j}\right\} - \frac{1}{2}\left[s(n) + s_\varepsilon(n)\gamma^2\sigma_\varepsilon^2 + s_\zeta(n)\sigma_\zeta^2\right]$$

5.3. Introducing learning

Though there are alternatives (e.g. Holmström and Tirole, 2001) we assume that the affine factor model of the term structure analyzed in Section 3 is a restricted (affine) version of the consumption-based asset pricing model, where the factors are to be interpreted as determinants of the intertemporal marginal rate of substitution. In this framework we can refer to Hansen and Singleton (1983) who prove that a necessary condition for asset returns to have predictable components is that (1) agents be risk averse and (2) the conditional mean of the growth rate of consumption varies over time. Since the empirical "evidence for predictability survives at reasonable if not overwhelming levels of statistical significance" (Campbell, 2000, p.9) we accept that these requirements have to be met by our theoretical model. This directly implies that at least one of the parameters $(A, \beta, \gamma, \delta, \sigma_\varepsilon^2)$ in the equation (46) must vary over time. Since β and γ characterize the utility function, we assume that from the consumer's point of view uncertainty is more important concerning the other parameters $(A, \delta, \sigma_\varepsilon^2)$, characterizing the production function. For the sake of simplicity we choose the conditional variance of the productivity shock σ_ε^2, but acknowledge that the other two parameters can also be subject to learning (or subjective probability adjustment in general) effects. This is

[8] Stock and Watson (2003) calculate the contemporaneous correlation between row series of US macroeconomic data and the four-quarter growth rate of GDP. They find that in the period 1984-2002 while the correlation coefficient for the 90-day T-bill rate is 0,39, the same indicator for price inflation (measured by the GDP deflator) is only 0,15 and for the 10-year T-bond rate is 0,02.

[9] If we used the quadratic approximation in the consumption equation, then the fourth moment of the productivity shock would also show up in the approximation.

clear also from the fact that the link between capital and output (the two variables, which can directly be measured) is the “joint product” of all the three parameters.

With the hypothesis of rational learning, the state of the economy presented in the model based on individual optimization depends on three factors: real consumption, inflation and the Jensen-factor. This setting is enough to allow for shocks coming from the real and nominal side of the economy as well as shocks coming from the financial markets. The yield curve becomes a joint product of the monetary authority (influencing the rate of inflation) and financial markets. At this level of generality many types of models (e.g. Markov-switching models or threshold VAR) could be used to describe the law of motion of the state variables, not just affine models as we do.

In order to make use of the stylized facts discovered by the empirical literature we have to calculate long term interest rates and holding period excess returns.

(55) $$R_{\infty,t} \equiv \lim_{n\to\infty} R_{n,t} = -\ln\beta + \gamma\overline{\Delta c} + \overline{\pi} - \frac{1}{2}\left[s(\infty) + s_{\varepsilon}(\infty)\gamma^2\sigma_{\varepsilon}^2 + s_{\zeta}(\infty)\sigma_{\zeta}^2\right]$$

According to Ang and Piazzesi (2003) very long term interest rates are overwhelmingly determined by latent variables. Many have noted that a shift in the conduct of monetary policy will likely lead to a change in the behavior of the term structure (e.g. Rudebusch (1995), Fuhrer (1996)). However, the results of Rudebusch and Wu (2004) suggest that the linkage is perhaps more subtle than is commonly appreciated. Stock and Watson (2003) also found that the direct effect of monetary policy rule changes on macroeconomic volatility is small. Moreover, since expected growth rate of consumption and inflation are most probably influenced by current and past values of consumption growth and inflation, which aren’t latent variables at all, our candidate for learning are again the conditional variance terms. Sensier and van Dijk (2003) test for a change in the volatility of 214 US macroeconomic time series over the period 1959-1999. They find that both consumption and inflation have experienced multiple breaks in unconditional volatility during this period. Even though the time series experienced a break in conditional mean, most of the reduction in volatility appears to be due to changes in conditional volatility. Volatility changes are more appropriately characterized as an instantaneous break rather than as a gradual change. As Muth (1960) already showed, in such an environment exponential weighting of past values (or constant gain learning) is the optimum way of forming expectations.

Since the Jensen-factor depends more on current and recent values of shocks, it will be correlated with contemporaneous or recent values of macroeconomic variables, but not completely, because it contains further information due to market participants’ subjective expectation about future course of monetary policy, technological progress and to possible Peso-problems as well. Correlation with macroeconomic variables can produce cyclical behavior of the Jensen-terms and hence of the term premia as it is documented in the empirical literature.

By substitution for the interest rates with different maturities from our closed form solution follows that the expected excess return only depends on some constant and the two conditional variances. This seems to be in contradiction with the empirical result of Tzavalis (2003) that the excess return (and hence its expected value as well) in the US yield curve can well be modeled by a one-dimensional AR(1) process. This might seem unexpected, since at least both consumption (real) and inflation (nominal) uncertainty should play a role in government bond prices, but as it has been shown by Giordani and Söderlind (2003, p 1046) for US professional forecasts for the period 1969-2001, though output growth uncertainty is slightly higher than inflation uncertainty, the two series are very strongly correlated (Corr=0,91). Of course, if the two series are not perfectly correlated, then the hypothesis of one-dimensional expected returns can be rejected, as it happened in Cochrane and Piazzesi (2003), but the 99% explanatory power reported by them is enough for us to follow this path.

5.4. Implications for economic modeling

The most direct implication of learning concerns the validity of the expectations hypothesis: market participants do use the weak form of the expectations hypothesis when ex ante forecasting future changes interest rates, but the constant term premium to be used in the EH has to be estimated from past data. In terms of macro-finance modeling this means that the term premium follows an AR(1) process but when more than one interest rates enters the model, their relation has to be modeled as if the term premium was constant.

Below we show four more pieces from the possible implications of learning about conditional volatilities.

Let us start with the quadratic approximation of the equilibrium rate of consumption growth, but substituting for the conditional variance originally assumed to be constant by its best estimate at time t:

$$E_t\{c_{t+1}\} = c_t + \frac{1}{\gamma}\boldsymbol{ln}[\beta(1-\delta+A)] + \xi_t$$

(56) where

$$\xi_t = \frac{1}{\gamma}\frac{1}{2}\frac{(1-\delta)A}{(1-\delta+A)^2}\hat{\sigma}^2_{\varepsilon,t}$$

In the case of rational learning best ex ante forecast of consumption is still the random walk with drift hypothesis, though ex post one might find partial predictability in the time series (which does not affect the unit root property).

If perceived uncertainty increases, then holding expected consumption constant current consumption will fall and aggregate savings will rise. This is empirically backed by Hahm and Steigerwald (1999), who study the effect of income uncertainty on consumption in a model that includes precautionary saving focusing on time-series variation in income uncertainty. Their time-series measure of income uncertainty is direct, since it is constructed from a panel of forecasts, where forecasters are asked not only about their point estimate but also about the distribution of their forecast (there are intervals of the form "income will increase between 2.0 and 2.9%" and the respondents assign probabilities to the intervals) They find evidence of precautionary saving in that increases in income uncertainty are related to increases in aggregate rates of saving. They also find evidence that anticipated income growth rates have less explanatory power for consumption growth rates after conditioning on income uncertainty. The evidence indicates the presence of forward-looking consumers who gradually adjust precautionary savings in response to changing income uncertainty.

Ex ante the best possible forecast of the Jensen-factor is its current value, thus its effect doesn't change with lengthening of the forecast horizon:

$$\textbf{(57)}\quad \frac{1}{n}E_t\{c_{t+n}-c_t\}=\frac{1}{\gamma}\boldsymbol{ln}[\beta(1-\delta+A)]+\xi_t \qquad \forall n\geq 1$$

Compared to the theoretical model the unit root property of the logarithm of consumption is still preserved but the previously constant drift is amended by a very small, but highly persistent AR(1) process, where the AR(1) process.

Ex post, however the whole story looks completely different. For longer time horizons we get the forecast by repeated substitution:

$$\textbf{(58)}\quad \frac{1}{n}E_t\{c_{t+n}-c_t\}=\frac{1}{\gamma}\boldsymbol{ln}[\beta(1-\delta+A)]+\bar{\xi}+\frac{1}{n}\frac{1-\rho^n}{1-\rho}(\xi_t-\bar{\xi})$$

where $\bar{\xi}$ is the unconditional mean of ξ_t.

From which follows that the effect of learning decays if the average rate of consumption growth is calculated over longer time-periods, since:

$$\textbf{(59)}\quad \frac{1}{n}E_t\{c_{t+n}-c_t\}=\frac{1}{\gamma}\boldsymbol{ln}[\beta(1-\delta+A)]+\bar{\xi}+\frac{1}{n}\frac{1-\rho^n}{1-\rho}(\xi_t-\bar{\xi})\xrightarrow{n\to\infty}\frac{1}{\gamma}\boldsymbol{ln}[\beta(1-\delta+A)]+\bar{\xi}$$

The second implication comes from the first order condition defining the log pricing kernel in our model:

$$\textbf{(60)}\quad R_{1,t}=-\boldsymbol{ln}\beta+\gamma E_t\{\Delta c_{t+1}\}+E_t\{\pi_{t+1}\}-\frac{1}{2}Var_t\{\gamma\Delta c_{t+1}+\pi_{t+1}\}$$

By rearranging the equation we get something similar to the expectational IS curve first derived by McCallum and Nelson (1997), but adjusted for risk:

$$\textbf{(61)}\quad \begin{aligned} c_t &= -\frac{1}{\gamma}\boldsymbol{ln}\beta+E_t\{c_{t+1}\}-\frac{1}{\gamma}(R_{1,t}-E_t\{\pi_{t+1}\})-\frac{\gamma}{2}Var_t\{\Delta c_{t+1}\}-\frac{1}{2\gamma}Var_t\{\pi_{t+1}\}= \\ &= -\frac{1}{\gamma}\boldsymbol{ln}\beta+E_t\{c_{t+1}\}-\frac{1}{\gamma}(R_{1,t}-E_t\{\pi_{t+1}\})-\frac{1}{2\gamma}\left[\left(\frac{A}{1-\delta+A}\right)^2\hat{\sigma}^2_{\varepsilon,t}+\hat{\sigma}^2_{\zeta,t}\right]\end{aligned}$$

Current consumption increases with expected level of consumption, decreases with forward looking real interest, and decreases with higher expected (perceived) level of risk. More risk averse agents (higher γ) care more about future consumption and care less about other variables, hence their consumption is closer to the pure random walk property.

The third consequence is the relation between yields on nominal and real investment. From the quadratic approximation of the growth rate of consumption (and suppressing the square of the conditional volatility) we can compare two ways of investment:

(62) $$R_{1,t} - E_t\{\pi_{t+1}\} + \frac{1}{2}\hat{\sigma}^2_{\zeta,t} \cong A - \delta + \frac{1}{2}\frac{(1-\delta)A - A^2}{(1-\delta+A)^2}\hat{\sigma}^2_{\varepsilon,t}$$

On the left hand side we sell our initial one unit of good, invest our money in government bond and see how much real goods we can buy for the proceeds at the end of the period (the standard definition of the real interest rate), while on the right hand side we invest our one unit of good in productive capital[10], have an output A and suffer a depreciation δ. In a deterministic world they should be equal, but in a stochastic world, both sides have to be adjusted for their risk: nominal investment for nominal risk and real investment for real risk. Though according to Giordani and Söderlind (2003) output uncertainty is slightly higher than inflation uncertainty, its multiplier is well below 1 for the usual values in empirical macro models, and so average return on real investment has to be higher than the real interest rate[11].

It might seem unexpected that higher nominal risk makes lower interest rate (both nominal and real) acceptable and higher real risk calls for higher interest rate. Standard CAPM would imply just the opposite, but here the variance is not that of the interest rate but of the pricing kernel, which moves in the opposite direction as the price of the corresponding nominal bond.

Our fourth interpretation concerns monetary policy reaction functions. Now we restate the closed form solution for the yield curve, but with the Jensen-factor substituted for the conditional variances of both consumption growth and inflation.

(63)

$$R_{n,t} = -\ln\beta + \gamma\overline{\Delta c} + \bar{\pi} + \gamma E_t\left\{\frac{1}{n}\sum_{j=1}^{n}\left(\Delta c_{t+j} - \overline{\Delta c}\right)\right\} + E_t\left\{\frac{1}{n}\sum_{j=1}^{n}\left(\pi_{t+j} - \bar{\pi}\right)\right\} + z_0(n) - z(n)(\xi_t - \bar{\xi})$$

---

[10] Here we use the approximation: $\boldsymbol{ln}(1-\delta+A) \cong A - \delta$

[11] It is possible that the two numbers are linked by the parameters of the model in the following way:

Output uncertainty (x) can be expressed as the sum of inflation uncertainty (y) and an other independent stochastic variable (z). According to Giordani and Söderlind (2003) the correlation between y and (y+z) is 0,91, which means, that the variance of z is about 20% of the variance of y. Long term interest rates are driven by both variances, but expected output variance is pre multiplied by the coefficient $\frac{(1-\delta)A - A^2}{(1-\delta+A)^2}$, which is about 0,1 for the usual values in AK-type growth models allowing for human capital. This means that the effect of z is reduced by a factor of 10 and the total effect on long term interest rate is determined by y (that is one single variable) up to (1-0,2/10)= 98%

where $z(n)$ is the time-invariant, but maturity dependent loading of the now demeaned AR(1) Jensen-factor and $z_0(n)$ is a maturity dependent constant. By taking unconditional expectations of both sides we get the unconditional mean of interest rates with different maturities:

(64) $\overline{R}_n = -\boldsymbol{ln}\beta + \gamma\overline{\Delta c} + \overline{\pi} + z_0(n)$

hence we need $z_0(n)$ to vary with the maturity in order to have an average slope of the yield curve different from zero (as it is required by the empirical evidence).

Specifically, the one period and the infinite maturity interest rates are

(65)
$$\begin{aligned} R_{1,t} &= \overline{R}_1 + \gamma E_t\{\Delta c_{t+1} - \overline{\Delta c}\} + E_t\{\pi_{t+1} - \overline{\pi}\} - z(1)(\xi_t - \overline{\xi}) \\ R_{\infty,t} &= \overline{R}_\infty - z(\infty)(\xi_t - \overline{\xi}) \end{aligned}$$

Our model doesn't allow for transactions cost, so the Dybvig-Ingersoll-Ross result about the impossibility of falling long term interest rates should hold. The reason, why it doesn't, is that the arbitrage opportunity claimed in their proof can only be used, if the numeric value of the lower bound of long term interest rates is known. If it is unknown, like in reality and in our proposal, then the maturity and composition of the arbitrage portfolio cannot be determined without uncertainty. This result is empirically confirmed by McCulloch and Kochin (2000). Put in another way, the DIR result in the case of learning only holds for the log pricing kernel, but since the market estimate of the unconditional variance (the infinite maturity limit of the conditional variances) changes over time, so must very long term interest rates do as well (in the opposite direction).

Substituting from the infinite maturity rate for the Jensen-factor in the short rate equation we get:

(66) $R_{1,t} = \overline{R}_1 + \gamma E_t\{\Delta c_{t+1} - \overline{\Delta c}\} + E_t\{\pi_{t+1} - \overline{\pi}\} + \dfrac{z(1)}{z(\infty)}\left[R_{\infty,t} - \overline{R}_\infty\right]$

In words: the short term interest rate has to deviate from its long run average if the growth rate of consumption, the rate of inflation or the long term interest rate deviate from their average. For risk neutral investors $\gamma = 0$ and hence the growth rate of consumption does not affect short term interest rates. This is compatible (without specifying the direction of causality) with a central bank, which doesn't have real variables (output, consumption or employment) in its reaction function (e.g. the European Central Bank) However, it is an interesting question for further research what happens if the degree of risk aversion of investors and of the central bank differs, as it is usually assumed about central bankers who hate inflation more than others do.

If we run a regression of the short term interest rate on just consumption growth and inflation, but without the long run interest rate, the error term will be highly serially correlated. This can lead to the illusion of interest rate smoothing in the monetary policy reaction function. Already Rudebusch (2002) raised this idea:

"Numerous studies have used quarterly data to estimate monetary policy rules or reaction functions that appear to exhibit a very slow partial adjustment of the policy interest rate. The conventional wisdom asserts that this gradual adjustment reflects a policy inertia or interest rate smoothing behavior by central banks. However, such quarterly monetary policy inertia would imply a large amount of forecastable variation in interest rates at horizons of more than three months, which is contradicted by evidence from the term structure of interest rates. The illusion of monetary policy inertia evident in the estimated policy rules likely reflects the persistent shocks that central banks face." (Abstract)

Our value added to Rudebusch's observation is twofold: (1) the persistent shocks he refers to come from the financial market in the form of time-varying estimates of the conditional variance of output and inflation and (2) the large amount of forecastable variation in interest rates based on ex post estimated policy rules can't show up ex ante for the very nature of rational learning.

A further consequence of a highly persistent Jensen-factor is the difficulty to prove the cointegration between short term interest rate and the rate of inflation, a key assumption behind the famous Taylor-rule. As Österholm (2003) says: "…there is very little support of cointegration between the included variables. This is a major shortcoming of the model since cointegration is a necessary condition for the relevance of the Taylor rule given the I(1) or near integrated behavior of the variables found in this study. This finding, together with the poor forecasting performance of the model, provides evidence that the rule is misspecified. Such misspecification might be one reason for the common finding in many papers that central banks seem to smooth the interest rate (for instance Clarida et al. (1998), (2000); introducing a serially correlated error term through misspecification is extremely likely to generate a significant parameter estimate on the lagged interest rate. This gives further support to the claim of Rudebusch (2002) and Söderlind et al (2003) that there is something wrong with the Taylor rule."

Our simple proposal is to amend the Taylor rule with the long term interest rate, but this brings us to some broader economic policy implications.

6. ECONOMIC POLICY IMPLICATIONS

In this section we explore some of the possible consequences for monetary policy.

6.1. The three goals of the Fed

Section 2A on monetary policy objectives of the Federal Reserve Act states that "The Board of Governors of the Federal Reserve System and the Federal Open Market Committee shall maintain long run growth of the monetary and credit aggregates commensurate with the economy's long run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates."[12] The first two

[12] [12 USC 225a. As added by act of November 16, 1977 (91 Stat. 1387) and amended by acts of October 27, 1978 (92 Stat. 1897); Aug. 23, 1988 (102 Stat. 1375); and Dec. 27, 2000 (114 Stat. 3028).]

of the three goals attracted already much attention, but the third one, moderate long term interest rates has been until recent times by and large completely forgotten. For example, the 1990 Economic Report of the President said, “My Administration will…support a credible, systematic monetary policy program that sustains maximum economic growth while controlling and reducing inflation.” Long term interest rates, which have a rather direct effect on the budget deficit due to the high level of US government debt, are not even referred to.

Federal Reserve Vice President Judd (1999) is absolutely explicit about the reason of the current practice: “The goals of ‘stable prices’ and ‘moderate long-term interest rates’ are related because nominal interest rates are boosted by a premium over real rates equal to expected future inflation. Thus, ‘stable prices’ will typically produce long-term interest rates that are ‘moderate’.”

The already more than ten years long line of research on Taylor-rules in the theory of monetary policy concentrates only on price stability measured by the rate of inflation and full employment measured by the output gap, without even mentioning long term interest rates as a possible argument in monetary policy rules. Another – though much thinner - line of research on monetary policy is the rule put forward by McCallum (1994) that allows for the slope of the yield curve that is the difference between long and short term interest rates.[13] Taylor-rules became much more fashionable in the last decade than McCallum-rules despite the evidence that since October 1987 the Fed increases the short term rate when there exists a widening of the spread signaling higher expected future inflation [See Tzavalis (2003), p.78] Nevertheless, even the McCallum-rule allows for long term interest rates in the monetary policy reaction function because of its signaling value about inflationary expectations and not because the Fed might have some preference over long term interest rates per se.

In our simplified consumption-based intertemporal asset pricing model the yield curve is a joint product of the real economy, monetary policy and financial markets represented by consumption, inflation and perception of future risk respectively. On the very long run (since the consumption to output ratio is strictly between 0 and 1) the rate of consumption growth equals the rate of output growth. If high rate of output growth and a low rate of inflation are defined as the first and second goal of the central bank, it is difficult to say anything specific about the desirable level of the long term interest rate. This might be the reason why the third goal of the Fed has almost always been forgotten. The adjective “moderate” reflects this problem. Nevertheless, long term interest rates already in the past and especially recently had an important effect on central banks’ monetary decisions. A possible interpretation of this fact is that if (and only if) the effect of output and inflation is accounted for, then long term interest rates can deliver information about aggregate perception of future risk. Since risk can be very costly in terms of social welfare in a world where many investors are risk averse it does make sense for the central bank to care about long term interest rates and hence long term interest rates can and have to be an independent argument in the monetary policy reaction function.

[13] McCallum (1994), Hsu and Kugler (1997) and Romhányi (2002) consider the effect of an AR(1) type yield spread entering the monetary policy reaction function on empirical tests of the EH and find it to be a possible explanation.

Recently there are signs of change with respect to the explicit role of long term interest rates in monetary policy. An interesting conversion of the relation between risk and long term interest rates is the non-traditional mean of monetary policy put forward by Bernanke (2002) member of the Board of Directors at the Fed, who proposed that central banks should try to keep long term bond prices above some specific level. He has suggested that short term interest rates being at the level or close to zero did not mean that the central bank could not drive long term bond rates down as long as the central bank announced that it would peg interest rates on long-term bonds at a very low interest rate (possibly zero) and stood ready to purchase any amounts of these bonds at this low rate. Another version of this same idea is of Orphanides and Wieland, (2000), who argue that in order to lower long-term bond rates the central bank has to convince the markets that it will continue to pursue a zero-interest-rate policy for a considerable time even after the deflation is over. Then, as is suggested by the expectations hypothesis of the term structure, because long-term bond rates are an average of the expected future short-term rates, long-term interest rates would necessarily fall. Indeed, this strategy is complementary to that of Bernanke (2002) because it is a way of committing to more expansionary policy in the future even after the economy has bounced back. The problem with this line of reasoning is that "the evidence that risk (term) premiums can be affected by changing the supply of long-term versus short-term bonds in the hands of the public is, unfortunately, far from clear" [Ito and Mishkin (2004), p.44]. The single historical example from the US in the 1960s has been generally viewed as a failure. This draws the attention to the clear distinction between compatibility and causality. Our model only tells what macro processes, market believes and asset prices are compatible with each other, but this doesn't tell anything about their causal relation. Nevertheless, central bank credibility (possibly backed by appropriate communication) is certainly an efficient way to decrease market perception of risk.

6.2. The Maastricht criterion on long term interest rate

On one hand one of the famous Maastricht criteria constituting the conditions of admissibility to the European Monetary Union is long term interest rates exceeding the reference value by not more than 2 percentage points, where the reference is the average of the three EU-member countries with the lowest rate of inflation.

If we accept the idea that the very long term interest rate can be interpreted as a measure of aggregate risk in the economy perceived by market participants, then the Maastricht criterion on convergence of long term interest rates is a natural requirement in an optimum currency area, where different regions must face the same shocks, otherwise the lack of exchange rate flexibility renders very costly the accommodation of idiosyncratic shocks.

Up to now only the other criteria (inflation, exchange rate stability, government deficit and government debt) caused problems to former candidate countries. May be this is the reason why the definition of the interest rate criterion is far the least elaborated. For the harmonized CPI and the government deficit and debt statistics there are whole volumes of manuals to specify the method of calculation while evaluation of the long term interest rate criterion is based simply on the yield to maturity of the longest available government security without specifying it any further.

On the other hand, when formulating the mission of the European Central Bank and since than there was much debate whether the ECB should aim full employment or solely concentrate on price stability. Nowhere has been raised the question of long term interest rates as a possible final goal.

However the question arises, whether interest rate smoothing on behalf of the central bank (that might be optimal if information about the state of the economy is noisy or available only with some lag) can lengthen the effect of short term interest rates on long term interest rates and hence only very long term interest rates will be free from the effect of interest rate smoothing. This might be an issue when pondering about the appropriateness of 10 year maturity coupon (not even zero coupon!) rates usually used to judge about this Maastricht criterion. From our analysis follows that the methodological problem of estimating infinite maturity interest rates should be addressed in order to give the interest rate criterion its true meaning.

6.3. Yield curve estimation with a new functional form

Measuring real world long term interest rates is far from being simple. Government bonds with a maturity above one year almost always pay coupon at discrete points over their lifespan implying that these bonds only can be treated as portfolios of different amounts of zero coupon bonds with different maturities. The estimated zero coupon yield curve is the result of a numeric optimization procedure. The optimum yield curve is the curve which has to be used to discount nominal cash flows in order to minimize pricing errors with respect to the actual prices at which government bonds are treaded on the market. There are several methods for the numeric optimization procedure. Shiller and McCulloch (1990) uses cubic spline approximation, which fits very well for medium term interest rates, but necessarily explodes at the long end of the yield curve, so it does not produce an approximation of the infinite maturity interest rate. (Most of the above mentioned empirical studies about the US yield curve use either McCulloch's calculations or some other data based on spline approximation.) Nelson and Siegel (1987) [further developed by Svensson(1994)] use a special combination of exponential function. This method almost surely converges to some finite value at the end of the yield curve, but might not fit at short maturities.

Both of these methods to estimate the yield curve only use contemporaneous market bond prices. Recently a new line of research (e.g. Diebold and Li, 2003 and De Rossi, 2004) raised the need and the possibility to use panel estimates of the yield curve, which use both cross section (over maturity) and longitudinal (over time) information. Based on our double affine model it is possible to calculate a panel estimate of the yield curve, where in the first step the vector of mutually orthogonal state variables at each point of time has to be estimated together with the autocovariance matrix, and in the second step, the vector of estimated state variables has to be regressed on macroeconomic data in order to find the economic and latent factors. An important advantage of the proposed functional form is beyond its simplicity and generality that it almost surely produces a finite value of the infinite maturity interest rate the development of which we are very interested in.

7. CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

We started in the introduction with three main inconsistencies between the traditional approach of economics and finance. They concerned the role of long term interest rates, financial markets and risk. The broad idea of our paper is that these three inconsistencies can be solved in a unified framework: from a macroeconomic perspective the relevant variable characterizing the stance of financial markets is market participants' perception of aggregate level of risk, which in turn can be measured by the level of long term interest rates, but the expectations hypothesis has to be relaxed to a hypothesis of learning, possibly with non-linearities and Peso-problems.

As we saw, even an extremely simplified macro-finance model of the term structure with learning can produce important phenomena well known from the empirical literature. The neoclassical framework with learning seems to be rich enough to capture many anomalies, but we have to be aware of the preconditions for neoclassical modeling: deep market, absence of informational asymmetries and low transactions costs. As we cited already above from Merton and Bodie (2004), all these to be true at least approximately, very long time and stability is needed so that appropriate institutions can evolve in order to compensate for anomalies due to non-neoclassical human behavior.

Despite the validity of the above mentioned preconditions in the case of U.S. government bond market, time to time structural changes hit the economy. It takes substantial time for market participants to learn even the first conditional moment of the relevant stochastic processes from observable variables. For higher moments, such as conditional volatility refinement by learning is even more time-consuming. Moreover, according to the historical evidence for the US economy, conditional variances change infrequently but rapidly. The best way for market participants to be always prepared for such changes is the method of exponential weighted average to estimate current value. This method is equivalent with a constant gain learning rule. A large part of the empirical literature about the US economy supports the learning hypothesis of the tem structure of interest rates.

If the learning hypothesis is true, then financial markets have to be explicitly allowed for in macro models. This is very much in the line of the rapidly growing literature about macro-finance modeling of the yield curve (e.g. Diebold et al. (2005) and Rudebusch and Wu (2004)) From an economic policy point of view, market perception about aggregate level of risk in the economy can be taken into account, if the monetary policy reaction function allows for the long term interest rate, as it is already required by the legal mandate of the Fed.

Finally, we mention a few of the possible directions for further research. From one side, the already voluminous structural VAR literature about identification of monetary policy shocks can now be amended explicitly by the long term interest rate as the substitute for the market perception of aggregate level of risk. From the other side, latent variable models can to identify the different shocks hitting the economy. Both results can be contrasted with e.g. narrative or survey based evidence about market shocks and monetary policy shocks. The time series properties of the infinite maturity interest rate can be explored and several interest rates can be included into macro models with explicit law of motion for the rate of inflation. In more complex models the complete dichotomy of nominal and real variables prevailing in our model can be relaxed, while the business cycle properties of the Jensen-factor have to be explored as well. In the field of asset pricing the effect of rational learning on the parameter estimates of the underlying stochastic processes can be explored. Casting the problem in a two-country framework exchange rate dynamics can be compared with the predictions of the standard uncovered interest parity and purchasing power parity relations.

References

Ang, Andrew and Geert Bekaert, 2000, Short Rate Nonlinearities and Regime Switches, Manuscript.

Ang, Andrew and Geert Bekaert, 2001, Regime Switches in Interest Rates, Manuscript.

Ang, Andrew and Monica Piazzesi, 2003, A No-Arbitrage Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables; Journal of Monetary Economics 50, 745-787.

Bams, Dennis and Christian C.P. Wolff, 2003, Risk Premia in the Term Structure of Interest Rates: a Panel Data Approach, Journal of International Financial Markets, Institutions and Money 13, 211-236.

Bansal, Ravi and Hao Zhou, 2002, Term Structure of Interest Rates with Regime Shifts, Journal of Finance 57/5, 1997-2043.

Bekaert, Geert and Robert, J. Hodrick and David A. Marshall, 2001, "Peso Problem" Explanations for Term Structure Anomalies, Journal of Monetary Economics.

Bernanke, Ben S., 2002, "Deflation: Making Sure "It" Doesn't Happen Here," remarks before the National Economists' Club, November 21st.

Bernanke, Ben S. and Alan S. Blinder, 1992, The Federal Funds Rate and the Channels of Monetary Transmission, American Economic Review, 901-921.

Bossaerts, Peter, 1998, On Rationality in Financial Markets, California Institute of Technology Working Paper.

Bossaerts, Peter, 2002, The Paradox of Asset Pricing, Princeton University Press.

Bullard, James and Kaushik Mitra, 2002, Learning about Monetary Policy Rules, Journal of Monetary Economics 49, 1105-1129.

Campbell, John Y., 1995, Some Lessons from the Yield Curve, Journal of Economic Perspectives 9/3, 129-152,

Campbell, John Y., 2000, Asset Pricing at the Millennium, Journal of Finance 55, 1515-1567.

Campbell, John Y. and John H. Cochrane, 1999, Explaining the Poor Performance of Consumption-Based Asset Pricing Models, NBER, No.7237.

Campbell, John Y. and Robert Shiller, 1991, Yield Spreads and Interest Rate Movements: A Bird's Eye View, Review of Economic Studies, 495-514.

Christiansen, Charlotte, 2003, Testing the Expectations Hypothesis Using Long-Maturity Forward Rates, Economics Letters 78, 175-180.

Cochrane, John, H., 2001, Asset Pricing, (Princeton University Press).

Cochrane, John H. and Monica Piazzesi, 2005, Bond Risk Premia, American Economic Review (forthcoming).

Collin-Dufresne, Pierre and Robert S. Goldstein and Christofer S. Jones, 2004, Can Interest Rate Volatility be Extracted from the Cross Section of Bond Yields? An Investigation of Unspanned Stochastic Volatility, NBER Working Paper No.10756.

De Rossi, Giuliano, 2004, Kalman Filtering of Consistent Forward Rate Curcves: a Tool to Estimate and Model Dynamically the Term Structure, Journal of Empirical Finance 11, 277-308.

Diebold, Francis X. and Canlin Li, 2003, Forecasting the Term Structure of Government Bond Yields, NBER Working Paper No. 10048.

Diebold, Francis X. and Monica Piazzesi and Glenn D. Rudebusch, 2005, Modeling Bond Yields in Finance and Macroeconomics, American Economic Review Papers and Proceedings.

Dotsey, Michael and Christofer Otrok, 1995, The Rational Expectations Hypothesis of the Term Structure, Monetary Policy, and Time-Varying Term Premia, Federal Reserve Bank of Richmond, Economic Quarterly, Volume 81/1, 65-81.

Duffie, Darrell, 1996, Dynamic Asset Pricing Theory (Princeton University Press).

Dybvig, Philip H., Ingersoll, Jonathan E. Jr. and Ross, Stephen A., 1996, "Long Forward and Zero-Coupon Rates Can Never Fall", The Journal of Business 69/1, 1-25.

Engsted, Tom and Carsten Tanggaard, 1994, Cointegration and the US Term Structure, Journal of Banking and Finance 18, 167-181.

Evans, George W. and Seppo Honkapohja, 2001, Learning and Expectations in Macroeconomics (Princeton University Press).

Evans, Martin D. D. and Karen K. Lewis, 1994, Do Stationary Risk Premia Explain It All?, Journal of Monetary Economics 33, 285-318.

Favero, Carlo A. and Federico Mosca, 2000, Uncertainty on Monetary Policy and the Expectations Model of the Term Structure of Interest Rates, Manuscript.

Foster, Dean P. and Dan B. Nelson, 1996, Continuous Record Asymptotics for Rolling Sample Variance Estimators, Econometrica 64/1, 139-174.

Fuhrer, Jeffrey C., 1996, Monetary Policy Shifts and Long-Term Interest Rates, Quarterly Journal of Economics, 1183-1208.

Gil-Alana, Luis A., 2004, Long Memory in the U.S. Interest Rate, International Review of Financial Analysis 13, 265-276.

Giordani, Paolo and Söderlind, Paul, 2003, “Inflation Forecast Uncertainty”, European Economic Review 47, 1037-1059.

Gregory, Allan W. and Graham M. Voss, 1991, The Term Structure of Interest Rates: Departures from Time-Separable Expected Utility, The Canadian Journal of Economics 24/4, 923-939.

Hahm, Joon-Ho and Douglas G. Steigerwald, 1999, Consumption Adjustment under Time-Varying Income Uncertainty, Review of Economics and Statistics 81(1), 32-40.

Hall, Robert E., 1978, Stochastic Implications of the Life Cycle - Permanent Income Hypothesis: Theory and Evidence, Journal of Political Economy 86, 971-987

Hansen, Lars Peter and Thomas J. Sargent, 1996, Recursive Linear Models of Dynamic Economies, Manuscript.

Hansen, Lars Peter and Kenneth J. Singleton, 1983, Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns, Journal of Political Economy 91/2, 249-265.

Hicks, J., 1939, Value and Capital (Oxford University Press).

Holmström, Bengt and Jean Tirole, 2001, LAPM: A Liquidity-Based Asset Pricing Model, Journal of Finance 56, 1837-1867.

Hsu, Chiente and Peter Kugler, 1997, The Revival of the Expectations Hypothesis of the US Term Structure of Interest Rates, Economics Letters 55, 115-120.

Ito, Takatoshi and Frederic S. Mishkin, 2004, Two Decades of Japanese Monetary Policy and the Deflation Problem, NBER, 10878.

James, Jessica and Nick Webber, 2000, Interest Rate Modelling, (John Wiley & Sons, Ltd).

Judd, John P., 1999, The Goals of U.S. Monetary Policy, Federal Reserve Bank of San Francisco Economic Letter, 99-04.

Jordà, Òscar. and Kevin D. Salyer, 2003, The Response of the Term Rates to Monetary Policy Uncertainty, Review of Economic Dynamics 6, 941-962.

Lanne, Markku, 1999a, Near Unit Roots and the Predictive Power of Yield Spreads for Changes in Long-Term Interest Rates, Review of Economics and Statistics 81/3, 393-398.

Lanne, Markku, 1999b, Testing the Expectations Hypothesis of the Term Structure of Interest Rates in the Presence of a Potential Regime Shift, Bank of Finland Discussion Paper, 20/99.

Lucas, Robert E. Jr., 1978, Asster Prices in an Exchange Economy, Econometrica 46/6, 1429-1445.

McCallum, Bennett T., 1994, Monetary Policy and the Term Structure of Interest Rates, NBER, No.4938.

McCallum, Bennett T. and Edward Nelson, 1997, An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis, NBER Working Paper No.5875.

McCulloch, J. Huston, 2000, Long Forward and Zero-Coupon Rates Indeed Can Never Fall, but are Indeterminate: a Comment on Dybvig, Ingersoll and Ross, Ohio State University, Working Paper #00-12.

McCulloch, J. Huston and Kochin, Levis A., 2000, "The Inflation Premium Implicit in the US Real and Nominal Term Structures of Interest Rates", Ohio State University, Working Paper #98-12 (Revised version, September 2000).

Merton, Robert C. and Zvi Bodie, 2004, The Design of Financial Systems: Towards a Synthesis of Function and Structure,NBER, No.10620.

Mulligan, Casey B., 2004, Robust Aggregate Implications of Stochastic Discount Factor Volatility, NBER Working Paper No. 10210.

Muth, John F., 1960, Optimal Properties of Exponentially Weighted Forecasts,Journal of the American Statistical Association, 299-306.

Nelson, Ch.R. and Siegel, A., 1987, „Parsimonious Modeling of Yield Curves", Journal of Business 60, 473-489.

Orphanides, Athanasios and Volker Wieland, 2000, Efficient Monetary Policy Design near Price Stability, Journal of Japanese and International Economics 14, 327-356.

Österholm, Pär, 2003, The Taylor Rule: A Spurious Regression?, Uppsala University Working Paper, No.20.

Österholm, Pär, 2004, Kiling Four Unit Root Birds in the US Economy with Three Panel Unit Root Test Stones, Applied Economic Letters 11, 213-216.

Romhányi, Balázs, 2002, Extending the McCallum Monetary Policy Reaction Function for More than Two Periods, Economics Letters 77, 227-232.

Rudebusch, Glen D., 1995, Federal Reserve Interest Rate Targeting, Rational Eexpectations and the Term Structure, Journal of Monetary Economics, 245-274.

Rudebusch, Glen D., 2002, Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia, Journal of Monetary Economics, 1161-1187.

Rudebusch, Glen D and Tao Wu, 2004, The Recent Shift in Term Structure Behavior from a No-Arbitrage Macro-Finance Perspective, Federal Reserve Bank of San Francisco, Working Paper Series, No.25.

Sensier, Marianne and Dick van Dijk, 2003, Testing for Volatility Changes in US Macroeconomic Time Series, Centre for Growht and Business Cycle Research, The University of Manchester, Discussion Paper Series, No.36.

Shiller, Robert J. and J. Huston McCulloch, 1990, The Term Structure of Interest Rates, in: B.M. Friedman and F.H. Hahn, eds., Handbook of Monetary Economics, Vol.1., (North-Holland, Amsterdam) 627-722.

Schich, Sebastian T., 2000, What the Yield Curve Say about Inflation:. Does it Change Over Time?, OECD Economics Department Working Paper, No. 227.

Smets, Frank and Kostas Tsatsaronis, 1997, Why does the Yield Curve Predict Economic Activity? Dissecting the Evidence for Germany and the United States, BIS Working Paper, No. 49.

Söderlind, Paul and Ulf Söderström and Anders Vredin, 2004, Dynamic Taylor Rules and the Predictability of Interest Rates, forthcoming in Macroeconomic Dynamics.

Stock, James H. and Mark W. Watson, 1989, New Indeces of Coincident and Leading Indicators, in: Olivier Blanchard and Stanley Fischer (eds.), NBER Macroeconomic Annual Vol.4.

Stock, James H. and Mark W. Watson, 2003, Has the Business Cycle Changed? Evidence and Explanations, in: Monetary Policy and Uncertainty, Federal Reserve Bank of Kansas City, 9-56.

Svensson, Lars E.O., 1994, Estimating and Interpreting Forward Interest Rates. Sweden 1992-1994, CEPR, 1051.

Taylor, John B., 1993, Discretion versus Policy Rules in Practice, Carnegie-Rochester Conference Series on Public Policy 39, 195-214.

Tzavalis, E., 2003, The Term Premium and the Puzzles of the Expectations Hypothesis of the Term Structure, Economic Modeling 21, 73-93.

Tzavalis, Elias and Michael R. Wickens, 1995, The Persistence in Volatility of the US Term Premium 1970-1986, Economics Letters 49, 381-389.

Tzavalis, Elias and Michael R. Wickens, 1997, Explaining the Failures of the Term Spread Models of the Rational Expectations Hypothesis of the Term Structure, Journal of Money, Credit, and Banking 29/3, 364-380.

Appendix A

Inference of the slope coefficients for the Campbel-Shiller regressions

Substituting the affine form introduced in (17) into (35):

(67)
$$\begin{aligned}&\{[g(n)+\boldsymbol{\gamma}'(n)\mathbf{z}_{t+k}]-[g(n)+\boldsymbol{\gamma}'(n)\mathbf{z}_t]\}=\\&=\alpha(n,k)+\beta(n,k)\frac{k}{n-k}\{[g(n)+\boldsymbol{\gamma}'(n)\mathbf{z}_t]-[g(k)+\boldsymbol{\gamma}'(k)\mathbf{z}_t]\}+\varepsilon_{t+k}\end{aligned}$$

Taking expectations conditional on information available at time *t* and equating terms gives us one condition:

(68) $$-\boldsymbol{\gamma}'(n)(\mathbf{I}-\mathbf{A}^k)\mathbf{z}_t=\hat{\beta}(n,k)\frac{k}{n-k}[\boldsymbol{\gamma}'(n)-\boldsymbol{\gamma}'(k)]\mathbf{z}_t$$

while substitution of the closed form solution derived in (19) gives a second condition

(69)
$$\begin{aligned}&-\left\{\boldsymbol{\delta}'(n)+[\boldsymbol{\gamma}'(1)-\boldsymbol{\delta}'(1)]\tfrac{1}{n}\sum_{j=0}^{n-1}\mathbf{A}^j\right\}(\mathbf{I}-\mathbf{A}^k)\mathbf{z}_t=\\&=\hat{\beta}(n,k)\frac{k}{n-k}\left\{\boldsymbol{\delta}'(n)-\boldsymbol{\delta}'(k)+[\boldsymbol{\gamma}'(1)-\boldsymbol{\delta}'(1)]\left(\tfrac{1}{n}\sum_{j=0}^{n-1}\mathbf{A}^j-\tfrac{1}{k}\sum_{j=0}^{k-1}\mathbf{A}^j\right)\right\}\mathbf{z}_t\end{aligned}$$

The stylized empirical result of Campbell and Shiller (1991) is that the slope coefficient is an asymptotically affine function of the maturity *n*:

(70) $$0<\lim_{n\to\infty}\left|\frac{\partial}{\partial n}\hat{\beta}(n,k)\right|<\infty$$

Inspection of the first condition shows that

(71) $$0<\lim_{n\to\infty}\|\boldsymbol{\gamma}(n)\|<\infty$$

otherwise the LHS of the equation would converge to zero or infinity, while the RHS to the finite value of $0<\left|-const*k\boldsymbol{\gamma}'(k)\mathbf{z}_t\right|<\infty$.

Before taking limit in the second condition in order to get the result for the Jensen-term we have to use the following calculation:

(72)
$$\begin{aligned}&\lim_{n\to\infty}\left(\tfrac{1}{n}\sum_{j=0}^{n-1}\mathbf{A}^j\right)(\mathbf{I}-\mathbf{A}^k)=\lim_{n\to\infty}\frac{1}{n}\left(\sum_{j=0}^{n-1}\mathbf{A}^j-\sum_{j=k}^{n+k-1}\mathbf{A}^j\right)=\lim_{n\to\infty}\frac{1}{n}\left(\sum_{j=0}^{k-1}\mathbf{A}^j-\sum_{j=n}^{n+k-1}\mathbf{A}^j\right)=\\&=\left(\sum_{j=0}^{k-1}\mathbf{A}^j\right)\lim_{n\to\infty}\frac{1}{n}(\mathbf{I}-\mathbf{A}^n)=\mathbf{0}\end{aligned}$$

Which is true iff none of the eigenvalues of the matrix A are outside the unit circle. From this follows that

(73) $$0 < \lim_{n\to\infty} \left\| \Omega(n) \right\| < \infty$$

Appendix B

Inference of zero coupon rates from the investor's problem

The problem of the investor is

$$U_t = E_t \sum_{j=0}^{\infty} \beta^j u_D\left(C_{t+j}\right) \rightarrow \boldsymbol{max}$$

s.t. $\forall j \geq 0$

(74) $$\Theta_{t+j} = \left\{ P_{t+j}\left[f\left(K_{t+j-1}\right) * \varepsilon_{t+j}\right] + \sum_{k=1}^{\infty} q_{k,t+j-k} \right\} -$$

$$- \left\{ P_{t+j}C_{t+j} + P_{t+j}\left[K_{t+j} - (1-\delta)K_{t+j-1}\right] + \sum_{k=1}^{\infty} p_{k,t+j}q_{k,t+j} \right\} = 0$$

The use of Lagrange multipliers in stochastic case is somewhat different from the deterministic case. [See e.g. Hansen and Sargent (1996)]

$$\boldsymbol{L}_t = U_t + E_t\left\{ \sum_{j=0}^{\infty} \beta^j \lambda_{t+j}\Theta_{t+j} \right\} = E_t \sum_{j=0}^{\infty} \beta^j \left\{ u\left(C_{t+j}\right) + \lambda_{t+j}\Theta_{t+j} \right\}$$

$$\frac{\partial \boldsymbol{L}_t}{\partial C_{t+j}} = E_t\left\{\beta^j\left[u'\left(C_{t+j}\right) - \lambda_{t+j}P_{t+j}\right]\right\} = 0 \quad \Rightarrow \quad \lambda_{t+j} = \frac{u_D'\left(C_{t+j}\right)}{P_{t+j}}$$

(75) $$\frac{\partial \boldsymbol{L}_t}{\partial K_{t+j}} = E_t\left\{\beta^{j+1}\lambda_{t+j+1}\left[P_{t+j+1}f_D'\left(K_{t+j}\right) * \varepsilon_{t+j+1} + P_{t+j+1}(1-\delta)\right] - \beta^j\lambda_{t+j}P_{t+j}\right\} = 0 \quad \Rightarrow$$

$$\Rightarrow \quad \beta\frac{\lambda_{t+j+1}}{\lambda_{t+j}} = \frac{P_{t+j}}{P_{t+j+1}\left[1 - \delta + f'\left(K_{t+j}\right) * \varepsilon_{t+j+1}\right]}$$

$$\frac{\partial \boldsymbol{L}_t}{\partial q_{n,t+j}} = E_t\left\{\beta^{j+n}\lambda_{t+j+n} - \beta^j\lambda_{t+j}p_{n,t+j}\right\} = 0 \quad \Rightarrow \quad \beta^n \frac{\lambda_{t+j+n}}{\lambda_{t+j}} = p_{n,t+j}$$

$$\frac{\partial \boldsymbol{L}_t}{\partial \lambda_{t+j}} = E_t\left\{\beta^j\Theta_{t+j}\right\} = 0 \quad \Rightarrow \quad \Theta_{t+j} = 0$$

Bond prices:

(76)

$$p_{n,t} = E_t\left\{\beta^n \frac{u'(C_{t+n})}{u'(C_t)}\frac{P_t}{P_{t+n}}\right\}$$

$$p_{1,t} = E_t\left\{\beta \frac{u'(C_{t+1})}{u'(C_t)}\frac{P_t}{P_{t+1}}\right\} = E_t\left\{\frac{P_t}{P_{t+1}\left[1-\delta+f'(K_t)*\varepsilon_{t+1}\right]}\right\}$$

Yields:

(77)

$$R_{n,t} = -\frac{1}{n}\boldsymbol{ln}\, p_{n,t} = -\frac{1}{n}\boldsymbol{ln}\, E_t\left\{\beta^n \frac{u'(C_{t+n})}{u'(C_t)}\frac{P_t}{P_{t+n}}\right\}$$

$$\beta\frac{u'(C_{t+1})}{u'(C_t)} = \frac{1}{1-\delta+f'(K_t)\varepsilon_{t+1}}$$

Table of Contents