Regensburger
DISKUSSIONSBEITRÄGE
zur Wirtschaftswissenschaft

University of Regensburg Working Papers in Business,
Economics and Management Information Systems

Risk and Policy Shocks on the US Term Structure

Enzo Weber[*], Jürgen Wolters[**]

March 16, 2010

Nr. 438

JEL Classification: E43, C32

Keywords: Expectations Hypothesis, Risk Premium, Policy Reaction Function, Persistence, Transitory Shocks

[*] Enzo Weber is Juniorprofessor of Economics at the Department of Economics and Econometrics at the University of Regensburg, 93040 Regensburg, Germany.
Phone: +49-941-943-1952, E-mail: enzo.weber@wiwi.uni-regensburg.de
[**] Jürgen Wolters is Professor of Econometrics at the Department of Economics at Free University Berlin, Boltzmannstr. 20, 14195 Berlin, Germany.
Phone: +49-30-838-52014, E-mail: juergen.wolters@fu-berlin.de

Risk and Policy Shocks on the US Term Structure[1]

Enzo Weber[2], Jürgen Wolters[3]

Abstract

We document two stylised facts of US short- and long-term interest rate data incompatible with the pure expectations hypothesis: Relatively slow adjustment to long-run relations and low contemporaneous correlation. We construct a small structural model which features three types of randomness: While a persistent monetary policy shock implies immediate identical reactions through the term structure, both a transitory policy shock and an autocorrelated risk premium allow for the sustained decoupling observed in the data. Indeed, we find important impacts and persistence of risk premia and a decomposition of policy shocks judging a larger part as transitory the longer the investment horizon.

Keywords: Expectations Hypothesis, Risk Premium, Policy Reaction Function, Persistence, Transitory Shocks
JEL classification: E43, C32

[1]We are grateful to participants of the 10th IWH-CIREQ Macroeconometric Workshop and the Quantitative Economics Colloquium at Freie Universität Berlin for their comments. Of course, all remaining errors are our own.

[2]Universität Regensburg, D-93040 Regensburg, Germany, enzo.weber@wiwi.uni-regensburg.de, phone: +49 (0)941 943-1952, fax: +49 (0)941 943-2735

[3]Freie Universität Berlin, Boltzmannstr. 20, D-14195 Berlin, Germany, juergen.wolters@fu-berlin.de, phone: +49 (0)30 838-52014, fax: +49 (0)30 838-54142

1 Introduction

The expectations hypothesis of the term structure (EHT) represents the most influential theory in explaining linkages between interest rates of different maturities. However, econometric research has notoriously failed in providing evidence for implications derived from the EHT (see e.g. Campbell and Shiller 1991, Hardouvelis 1994, Campbell 1995). Theoretically, the spread between the long- and short-term rate should have predictive power for future changes of both the short- and long-term rates.

In detail, a positive relation between future changes of short-term rates and the spread exists with a coefficient equal to the maturity of the respective long-term rate. This positive relation is due to the fact that according to EHT, the long-term rate arises as an average of expected future short-term rates. Therefore, an increase in the long-term rate relative to the short rate implies rising future short-term rates. Most econometric investigations do find a significant positive relation but with a coefficient far below what is expected theoretically. Moreover, there is empirical evidence (see e.g. Campbell and Shiller 1991, Hardouvelis 1994) that the spread is negatively related to future changes in the long-term interest rate, whereas it is theoretically expected that an increase in the current long rate relative to the current short rate is followed by a positive change in the next period.

Most empirical investigations of the EHT are restricted to assessing the implications either for the short-term rates or the long-term rates. However, analysing the implications for both rates simultaneously is econometrically more efficient. This holds especially when interest rates are non-stationary in the sense that their first differences are stationary, i.e. they follow processes integrated of order one (I(1)). In this case, in a first step one has to test whether the spread is stationary, meaning that long- and short-term rates have to be cointegrated with cointegration vector $(1, -1)$. According to the Granger representation theorem (see Engle and Granger 1987), a vector error correction model (VECM) exists with the spread as long-run relation. A sufficient condition for stability of such a system is that both variables adjust towards the equilibrium, implying a positive adjustment coefficient for the short-term rate and a negative one for the long-term rate. Remarkably, this negative relation, which is found empirically, is at odds with the pure EHT.

As a reaction to these empirical difficulties, the pure EHT has been extended by several amendments. Hamilton(1988) as well as Sola and Driffill(1994) allow for stochastic changes in regimes. Another approach is to take into account a time-dependent risk premium, which is often modelled by processes with autoregressive conditional heteroscedasticity, see e.g. Engle et al. (1987) or Tzavalis and Wickens (1995). Caporale and Caporale (2008) explain shifts in premia by political risk. A strand of literature including Dai and Singleton (2000), Duffee (2002) and Ang and Piazzesi (2003) models the dynamics of the yield curve by observed and unobserved factors. Another term structure model, which leads to analysing both rates simultaneously, is constructed by McCallum (2005), who combines autocorrelated risk premia with a monetary policy feedback rule including interest rate smoothing. Kugler (1997) generalised McCallum's result for any maturity, presenting equations for the short rate and the spread. Weber and Wolters (2009) derived the vector error correction (VEC) form of the model and tested its statistical properties for the whole US term structure.

This line of research significantly contributed to an economic explanation of the empirically observable deviations from the pure EHT. However, the underlying statistical models were not fully structural in the sense that they neglected implications of the theoretical model framework for the covariance matrix of the disturbances. In the present paper, we derive these implications, impose them in the VECM estimation and show that the constraints cannot be supported by the data.

We argue that these rejections are due to an overly restrictive feature implicit in the employed term structure models: In particular, policy shocks to the short rate are treated as fully persistent. It follows that via the averaging of (expected) future short rates in the EHT equation, long rates must be hit instantaneously with identical strength. While such a high cross-correlation implied by the model is not supported by the data, neither can it be empirically justified to treat the short rate as non-persistent (stationary). We resolve this dilemma by including a transitory shock in the monetary policy equation, in addition to the persistent one considered in the existing literature. Thereby, the mean reversion of the transitory part allows for a wedge in the tight EHT linkage. The relative contributions of the two types of shocks can be freely estimated from term structure data. This clear-cut separation represents a particular advantage of the proposed approach, since impacts and variance shares can be uniquely traced back to each of the structural innovations.

In addition to the risk premium and the persistent policy innovations, the transitory shock introduces a third source of randomness into a system of only two interest rates. This induces non-trivial problems in identifying and estimating the model. For identification purposes, we exploit the heteroscedasticity caused by the shift in variances of interest rates in the early 1980s after the Volcker disinflation period. Estimation is performed through Maximum Likelihood (ML) in the approximate structural VECM derived from the theoretical model equations. Indeed, we find considerable empirical support for the presence of non-persistent policy shocks in the US term structure. The effect gains importance when measured against longer-term bond yields, decoupling from the one-month rate. Furthermore, the estimates reveal sizeable risk premium effects, featuring rising persistence for higher maturities.

The reader can expect the following: The next section restates the model from Weber and Wolters (2009) and demonstrates the empirical rejection of the structural form. In section 3, we formulate the model including the transitory shock and develop its VEC representation. We conduct statistical inference on the parameters and investigate the relative importance of the different shocks. The last section concludes.

2 Persistent *or* Transitory Policy?

Herein, we shortly restate the model from Weber and Wolters (2009), based on McCallum (2005). This prepares the ground for extending the framework in the following section. To begin with, the pure EHT equates the return of a single n-period fixed-interest investment and the overall expected return of a series of n successive one-period investments. Combined with a maturity-dependent risk premium $v_t^{(n)}$, which is modelled as an autoregressive (AR) process of order one[4], the linearised form can be written as

$$R_t^{(n)} = \frac{1}{n}\sum_{i=0}^{n-1} E_t r_{t+i} + v_t^{(n)} \quad \text{with} \quad v_t^{(n)} = c_n + \rho_n v_{t-1}^{(n)} + u_t^{(n)} \ , \ 0 < \rho_n \leq 1 \ . \tag{1}$$

In (1), $R_t^{(n)}$ denotes the interest rate of maturity $n > 1$ and r_t of maturity one. The operator E_t stands for the expectation given all information available at time t. Given risk averse agents, long-term interest rates do contain term premia, in addition to information on future short rates; compare e.g. Tzavalis and Wickens (1998), who found stationary premia for US data, and Evans and Lewis (1994), where evidence for non-stationarity is established. We model time-dependent risk premia by an autoregressive process, which can capture both cases.

[4]Generalisations to higher-order processes are straightforward. In case of a higher-order AR process for the risk premium the structure of the adjustment parameters in (3) will remain the same, where ρ now measures the persistence of the premium as the sum of the AR coefficients. Additionally, this would imply further short-run dynamics, a fact that will be taken into account in our empirical model (4).

The second model equation introduces the stylised monetary policy rule

$$r_t = r_{t-1} + \lambda_n (R_t^{(n)} - r_t) + \varepsilon_t^{(n)} \; . \tag{2}$$

According to (2), the central bank conducts interest rate smoothing (with coefficient 1 as in McCallum 2005), but changes the short rate in response to the yield spread with feedback intensity $\lambda_n > 0$. Thereby, the spread may act as an indicator for monetary policy expansiveness (Laurent 1988), future economic growth (e.g. Estrella and Hardouvelis 1991) and expected inflation rates (Mishkin 1990), approximating forward-looking counter-cyclical policy behaviour; see also Johnson (1988) and McCallum (2005) in this context. The policy shock $\varepsilon_t^{(n)}$ is uncorrelated with the risk premium shock $u_t^{(n)}$.

As solution of equations (1) and (2), the VECM representation (up to constant terms) follows as[5]

$$\begin{pmatrix} \Delta R_t^{(n)} \\ \Delta r_t \end{pmatrix} = \begin{pmatrix} \lambda_n \rho_n + \rho_n - 1 \\ \lambda_n \rho_n \end{pmatrix} (R_{t-1}^{(n)} - r_{t-1}) + \begin{pmatrix} 1 & (1+\lambda_n)\theta_n \\ 1 & \lambda_n \theta_n \end{pmatrix} \begin{pmatrix} \varepsilon_t^{(n)} \\ u_t^{(n)} \end{pmatrix} \; , \tag{3}$$

where $\theta_n = n/(n - \lambda_n \sum_{j=1}^{n-1} (n-j)\rho_n^j)$. From (3) it follows that the change in the short-term rate is always positively related to the last period's spread, but with a coefficient far below maturity n. Normally, also the change of the long-term rate is positively related to the spread. Only for small (positive) values of λ_n and ρ_n this relation may become negative.

The corresponding reduced-form VECM of general lag length p is given by

$$\begin{pmatrix} \Delta R_t^{(n)} \\ \Delta r_t \end{pmatrix} = \begin{pmatrix} \alpha_{n1} \\ \alpha_{n2} \end{pmatrix} (R_{t-1}^{(n)} - \beta_{n2} r_{t-1} - c_n) + \sum_{i=1}^{p} \Gamma_{ni} \begin{pmatrix} \Delta R_{t-i}^{(n)} \\ \Delta r_{t-i} \end{pmatrix} + \begin{pmatrix} w_{1t}^{(n)} \\ w_{2t}^{(n)} \end{pmatrix} \; , \tag{4}$$

with adjustment vector $\alpha_n = (\alpha_{n1}, \alpha_{n2})'$, cointegration vector $\beta_n = (1, -\beta_{n2})'$, vector of white-noise errors $w_t^{(n)}$, 2×2 parameter matrices Γ_{ni} in the short-run dynamics and a constant constrained to the cointegrating term. Without imposing the restrictions from the impact matrix of the shocks, the structural system is exactly identified for $\beta_{n2} = 1$ in the sense that there is a one-to-one relationship between $\{\rho_n, \lambda_n\}$ and $\{\alpha_{n1}, \alpha_{n2}\}$. Nevertheless, besides determining the adjustment coefficients, the theoretical model implies a

[5] This is easily derived from equations (16) and (17) in the Appendix.

specific structure for the disturbances. In particular, if (3) is a valid representation of the data, it holds that

$$Cov\begin{pmatrix} w_{1t}^{(n)} \\ w_{2t}^{(n)} \end{pmatrix} = \begin{pmatrix} 1 & (1+\lambda_n)\theta_n \\ 1 & \lambda_n\theta_n \end{pmatrix} Cov\begin{pmatrix} \varepsilon_t^{(n)} \\ u_t^{(n)} \end{pmatrix} \begin{pmatrix} 1 & (1+\lambda_n)\theta_n \\ 1 & \lambda_n\theta_n \end{pmatrix}' . \qquad (5)$$

Given uncorrelatedness as the usual and sensible assumption for structural shocks, the covariance matrix on the right hand side of (5) includes only *two* unknowns, i.e. the variances, beyond ρ_n and λ_n that can already be recovered from the adjustment coefficients. Logically, this implies one restriction on the left-hand-side reduced-form covariance matrix, which contains *three* distinct entries. It follows that the theoretical model can be tested by checking the validity of the overidentifying restriction. Doing so, we estimate the VECMs by Maximum Likelihood (ML) both with and without the covariance restrictions from (5) (in addition to the α_n restrictions and $\beta_{n2} = 1$) and compute the Likelihood Ratio (LR) test statistics.

Our data set consists of monthly observations of US certificate of deposit rates ($n = 1, 3, 6$ months) and constant maturity bond yields ($n = 12, 24, 36, 60, 84, 120$ months) obtained from the Fed. The sample begins in 1979(8)[6] with the Volcker chairmanship and ends in 2008(12). Figure 1 shows the development of the 1-month, 3-year and 10-year interest rates, giving an impression of the dynamics in the US term structure.

The time series exhibit highly persistent behaviour. This is formally tested with an augmented Dickey-Fuller (ADF) test, where the null hypothesis is given by a unit root, as well as with a KPSS test, where the null hypothesis is switched to stationarity.[7] As can be seen from Table 1 we get unambiguous results from the two different tests: While ADF tests do not reject the null of non-stationarity, KPSS tests overwhelmingly reject the null of stationarity. Since the first differences of all interest rates are clearly stationary, all interest rates are I(1).[8]

Therefore, as Weber and Wolters (2009) point out, a necessary precondition for the validity of the structural specification (1), (2) is given by stationarity of the spreads. This would be achieved by pair-wise cointegration of the I(1) interest rates with restricted cointegration

[6] Up to 12 pre-sample values are used to estimate dynamic models.

[7] For performing ADF and KPSS tests as well as cointegration analyses see e.g. Kirchgässner and Wolters (2007, chapters 5 and 6).

[8] Note that the outcome of unit root tests and cointegration analysis is not an invariant property of the underlying variable but strongly depends on the sample period and the frequency of the observations; see e.g. Juselius (1999).

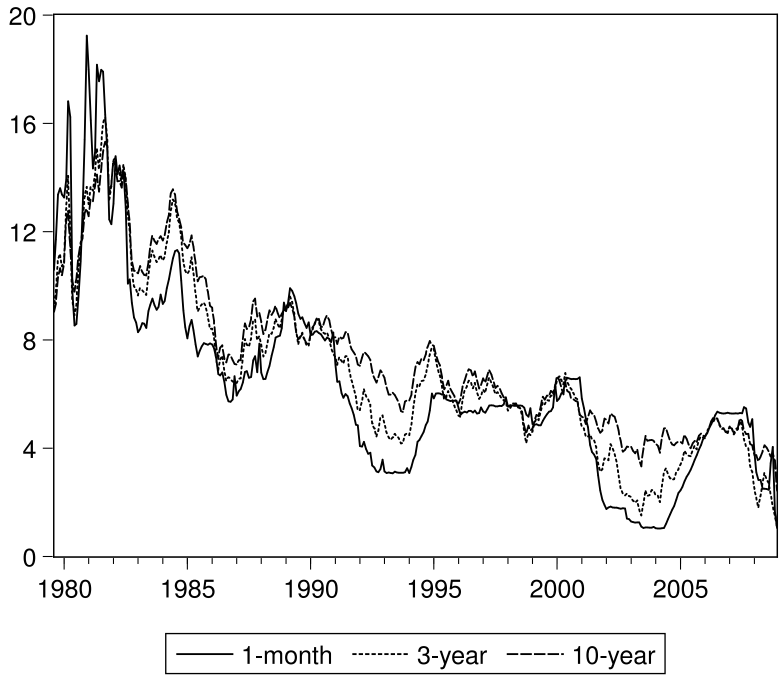

Figure 1: Interest rates

vector equal to $(1, -1)$. The first issue implies a rank of $r = 1$ of the long-run matrix in a bivariate non-stationary vector autoregression (VAR). Consequently, we test the null hypothesis of $r = 0$ (no cointegration) against the alternative of $r \geq 1$ (at least one cointegrating relation). Johansen (1995) trace tests uniformly reject the null hypotheses at least at the 3% significance level. The second issue points at the null hypothesis of $\beta_{n2} = 1$. LR tests in no case reject this parameter restriction. Thus, we are ready to check the structural covariance constraints as described above. The according LR test results are shown in Table 2.

All test statistics exceed the 1% $\chi^2(1)$ critical value of 6.64, leading us to reject the

n	1	3	6	12	24	36	60	84	120
ADF	-1.85	-2.14	-1.40	-1.03	-0.90	-0.96	-0.81	-0.70	-0.69
lags	9	16	6	6	6	2	2	2	2
KPSS	1.60	1.60	1.61	1.73	1.84	1.89	1.95	1.99	2.01

Notes: All ADF tests insignificant at any usual level; all KPSS tests significant at 1%; all models with constant; ADF lag length determined by Schwarz criterion; KPSS with Bartlett kernel and Newey-West bandwith selection (15 $\forall$ n); first differences clearly stationary.

Table 1: Unit root (ADF) and stationarity (KPSS) tests

n	3	6	12	24	36	60	84	120
lags	1-2,4,6-8	1-2,4,6-8, 11-12	1-2,7,11		1-2,9,11			1,6,9,11
LR	9.29	35.99	155.95	229.65	251.91	321.25	364.40	423.13

Table 2: LR tests on structural covariance restrictions

constraints. Thus, the theoretical model seems to be too restrictive when tested against US term structure data especially for longer-term bonds. To illustrate the underlying argument, note that in (3), the monetary policy shock $\varepsilon_t^{(n)}$ is bound to hit both interest rates with identical strength of 1. The reason is that every shock to the short rate that persists through time (as can be seen from (11) in the Appendix, ignoring $q_t^{(n)}$) is immediately incorporated in full into the long-rate via the EHT equation (1). Empirically however, such a high residual cross-correlation is not supported by the reduced-form estimates. Since this empirical cross-correlation between long and one-period rate shocks shrinks for longer maturities, evidence against the null hypothesis becomes harsher.

One possible solution would be to make the policy shock in (2) non-persistent (i.e., transitory), in order to avoid the full impact on expectations of all future short rates. However, this would contradict wide-spread consensus that interest rates are well approximated by I(1)-processes (see Mankiw and Miron 1986). Indeed, as shown above, neither were we able to reject the null hypothesis of non-stationarity of the one-month rate with an ADF test, nor to retain the reversed null hypothesis of stationarity with a KPSS test. That is, a more sophisticated modelling approach is called for, taking into account all relevant data properties.

3 Persistent *and* Transitory Policy

While the model from the previous section succeeded in explaining the adjustment of the interest rates to their spread, it clearly failed to provide an accurate description of the contemporaneous covariance. From the above discussion it can be seen that an appropriate model must feature both, interest rate persistence and mean reverting components in the policy shock. In order to achieve a clear-cut decomposition of these components, we include an additional transitory shock $z_t^{(n)}$ orthogonal to $u_t^{(n)}$ and $\varepsilon_{1t}^{(n)}$ in the policy rule (2):[9]

[9] We stick to the original notation even though the equation is changed by the new shock.

$$r_t = r_{t-1} + \lambda_n(R_t^{(n)} - r_t) + \varepsilon_{1t}^{(n)} + z_t^{(n)} \ . \tag{6}$$

For rendering the additional policy shock truly transitory, its long-run impulse response effect on the short rate must be constrained to zero, considering the whole system interaction. This can be ensured by specifying $z_t^{(n)}$ as the first-order moving average (MA) process $z_t^{(n)} = q_t^{(n)} - b_n q_{t-1}^{(n)}$, with $q_t^{(n)}$ white noise. Then, after the model solution (see the Appendix), the coefficient b_n can be assigned the value that meets the zero long-run constraint. Evidently, in this pure MA structure the initial impact of $q_t^{(n)}$ would fully cancel out in the following period, restricting the effect of the transitory shock to a single period. To allow for more general and prolonged impulse responses, we let $q_t^{(n)}$ follow the AR(1)-process $q_t^{(n)} = a_n q_{t-1}^{(n)} + \varepsilon_{2t}^{(n)}$, resulting in ARMA(1,1) dynamics in $z_t^{(n)}$.

From the system (1), (6), we derive the rational expectations solution (see the Appendix for details). The VEC form results as

$$\begin{aligned}
\begin{pmatrix} \Delta R_t^{(n)} \\ \Delta r_t \end{pmatrix} &= \begin{pmatrix} (\lambda_n\rho_n + \rho_n - 1)(1-a_n) \\ \lambda_n\rho_n(1-a_n) \end{pmatrix} (R_{t-1}^{(n)} - r_{t-1}) \\
&+ \begin{pmatrix} a_n\rho_n(1+\lambda_n) & a_n(1-\rho_n-\lambda_n\rho_n) \\ a_n\lambda_n\rho_n & a_n(1-\lambda_n\rho_n) \end{pmatrix} \begin{pmatrix} \Delta R_{t-1}^{(n)} \\ \Delta r_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & \frac{n+S(a_n)(a_n-b_n-\lambda_n b_n)}{n-\lambda_n a_n S(a_n)} & \frac{n(1+\lambda_n)}{n-\lambda_n\rho_n S(\rho_n)} \\ 1 & \frac{n-\lambda_n b_n S(a_n)}{n-\lambda_n a_n S(a_n)} & \frac{\lambda_n n}{n-\lambda_n\rho_n S(\rho_n)} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t}^{(n)} \\ \varepsilon_{2t}^{(n)} \\ u_t^{(n)} \end{pmatrix} \\
&+ \begin{pmatrix} -a_n & \frac{S(a_n)(\lambda_n a_n b_n - \rho_n(a_n-b_n)(1+\lambda_n)) - b_n n}{n-\lambda_n a_n S(a_n)} & \frac{-a_n n(1+\lambda_n)}{n-\lambda_n\rho_n S(\rho_n)} \\ -a_n & \frac{\lambda_n S(a_n)(a_n b_n - \rho_n(a_n-b_n)) - b_n n}{n-\lambda_n a_n S(a_n)} & \frac{-a_n\lambda_n n}{n-\lambda_n\rho_n S(\rho_n)} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1}^{(n)} \\ \varepsilon_{2t-1}^{(n)} \\ u_{t-1}^{(n)} \end{pmatrix} ,
\end{aligned} \tag{7}$$

with $S(x) := \sum_{i=1}^{n-1}(n-i)x^{i-1}$. Formally, (7) is in VARMA form. Since it is well known that invertible VARMA models can be approximated by higher-order VARs, we will estimate structural VECMs of sufficient order p, as given in reduced form by (4).

Within this empirical setup, we must ensure that the second policy shock $\varepsilon_{2t}^{(n)}$ is indeed transitory, i.e. that the according impulse responses of the short rate die out in the long run. Johansen (1995) shows that the matrix Ξ of long-run impacts of the reduced-form VECM errors on the endogenous variables takes the form

$$\Xi_n = \beta_{n\perp}(\alpha'_{n\perp}(I - \sum_{i=1}^{p}\Gamma_{ni})\beta_{n\perp})^{-1}\alpha'_{n\perp} \ , \tag{8}$$

with $\perp$ denoting the orthogonal complement (thus $\delta'\delta_\perp = 0$, where both δ and $\delta_\perp$ have full column rank). In our term structure model, the reduced-form residuals $w_t^{(n)}$ are connected to the fundamental shocks $e_t^{(n)}$ as in (7) by

$$\underbrace{\begin{pmatrix} w_{1t}^{(n)} \\ w_{2t}^{(n)} \end{pmatrix}}_{w_t^{(n)}} = \underbrace{\begin{pmatrix} 1 & \frac{n+S(a_n)(a_n-b_n-\lambda_n b_n)}{n-\lambda_n a_n S(a_n)} & \frac{n(1+\lambda_n)}{n-\lambda_n \rho_n S(\rho_n)} \\ 1 & \frac{n-\lambda_n b_n S(a_n)}{n-\lambda_n a_n S(a_n)} & \frac{\lambda_n n}{n-\lambda_n \rho_n S(\rho_n)} \end{pmatrix}}_{B_n} \underbrace{\begin{pmatrix} \varepsilon_{1t}^{(n)} \\ \varepsilon_{2t}^{(n)} \\ u_t^{(n)} \end{pmatrix}}_{e_t^{(n)}} . \tag{9}$$

Consequently, the long-run impacts of the structural innovations result as $\Xi_n B_n$. Specifying $\varepsilon_{2t}^{(n)}$ as a transitory shock to the policy rate thus amounts to imposing a zero restriction on the (2,2)-element of $\Xi_n B_n$. In particular, in our empirical work, this parallels determining b_n in the theoretical model as to ensure the above-discussed zero long-run influence on r_t.

For the model to be identified, the reduced form must deliver enough information for recovering all structural coefficients. In particular, (7) contains the seven unknown magnitudes λ_n, ρ_n, a_n and b_n as well as three variances of the shocks. From a reduced-form VECM as in (4) with $\beta_{n2} = 1$, two adjustment coefficients in α_n, two residual variances and one covariance can be obtained. Furthermore, one restriction is set on $\Xi_n B_n$. In sum, there is a lack of one piece of information.

We address this problem as follows: In general, the reduced-form covariance matrix acts as a source of information, since it merges the structural impact coefficients and variances in the quadratic form $Cov(w_t^{(n)}) = B_n Cov(e_t^{(n)}) B_n'$ following from (9). Now, suppose a break in structural variances at a specific time in the sample, such that $Cov(e_t^{(n)}|t \in R_1) \neq Cov(e_t^{(n)}|t \in R_2)$; R_i denotes the set of points in time belonging to the ith regime. Then, an extra reduced-form covariance matrix $Cov(w_t^{(n)}|t \in R_2)$ for the second sub-period can be obtained *in addition* to $Cov(w_t^{(n)}|t \in R_1)$ from the first sub-period. Therefore, the amount of available information would be doubled. Similar identification principles have been described for instance in Sentana and Fiorentini (2001) and Rigobon (2003).

Reconsidering Figure 1, it is straightforward to date a volatility shift to the early 1980s, coinciding with the end of the Volcker disinflation period. Accordingly, we specify breaks in the variances of both $\varepsilon_{1t}^{(n)}$ and $\varepsilon_{2t}^{(n)}$ in 1982(10). Since it is common sense that the high volatility of interest rates in the first years of the sample was triggered by monetary policy acting, the homoscedasticity assumption is kept for the remaining risk premium shock $u_t^{(n)}$. Thus, we introduce only two additional unknowns as the first two entries on the diagonal of $Cov(e_t^{(n)}|t \in R_2)$, but gain three pieces of information from the second-period reduced-form covariance matrix. Thereby, we achieve exact identification of our term structure model. Economically, while we allow for breaks in the variances of the policy shocks, the systematic part of the reaction function is assumed to remain unchanged. This specification is based on statistical evidence that shifts in the policy parameter λ_n proved to be clearly insignificant throughout all models. It seems that even though the Volcker disinflation generated considerably higher interest rate volatility, the fundamental reaction to the current spread does not differ measurably between the regimes. In that latter sense, the use of money and interest targets could be considered as equivalent, see Poole (1970).

Estimation is again performed by ML. For this purpose, we assume that the innovations are normally distributed within both regimes. This means that unconditionally, a normal mixture distribution applies. The likelihood function is constructed for the approximate structural VECM(p), taking into account all restrictions on the adjustment coefficients, the covariance matrix and the long-run impact matrix. We apply the numerical Newton algorithm to maximise the likelihood over the free parameters λ_n, ρ_n, a_n, b_n, the shock variances in both regimes, the constant restricted to the long-run relation with $\beta_{n2} = 1$ and the parameters from the short-run dynamics.

The lag lengths of the reduced-form models are the same as in Table 2. The large number of lags should appropriately approximate the VARMA dynamics from (7). When estimating the model, we encounter that the AR-coefficients a_n are generally insignificant, as can be seen from their t-values in Table 3.

n	3	6	12	24	36	60	84	120
t_n	0.31	0.24	-0.60	-0.82	-0.17	0.54	0.75	-0.78

Table 3: t-values for a_n

It follows that the transitory policy shock is virtually restricted to a duration of one period. In order to simplify the model and enhance efficiency, we impose $a_n = 0$. The parameter estimates with their standard errors in parentheses for the structural VECM, equations (16) and (17)) in the Appendix, are presented in Table 4. This system has the same conditional expectation as the system in (3), but a different structure in the residuals. Columns 2 to 5 in Table 4 show the coefficients estimated from the whole sample period 1979(8) to 2008(12). The variances of the permanent and transitory policy innovations for the first regime, 1979(8) to 1982(10), are given in columns 6 and 7 and for the second regime, 1982(11) to 2008(12), in columns 8 and 9.

Regime	I + II				I		II	
n	$\hat{\lambda}_n$	$\hat{\rho}_n$	$\hat{b}_n$	$\hat{\sigma}^2_u$	$\hat{\sigma}^2_{\varepsilon_1}$	$\hat{\sigma}^2_{\varepsilon_2}$	$\hat{\sigma}^2_{\varepsilon_1}$	$\hat{\sigma}^2_{\varepsilon_2}$
3	0.373 (0.149)	0.736 (0.055)	1.062 (0.363)	0.002 (0.001)	3.076 (0.712)	0.184 (0.129)	0.069 (0.009)	0.024 (0.011)
6	0.203 (0.063)	0.837 (0.036)	0.967 (0.222)	0.004 (0.001)	2.539 (0.594)	0.869 (0.447)	0.061 (0.007)	0.019 (0.006)
12	0.182 (0.040)	0.868 (0.028)	0.790 (0.094)	0.005 (0.002)	1.547 (0.370)	1.771 (0.581)	0.021 (0.010)	0.076 (0.013)
24	0.092 (0.026)	0.924 (0.021)	0.865 (0.103)	0.006 (0.003)	1.020 (0.250)	1.870 (0.607)	0.024 (0.010)	0.073 (0.013)
36	0.079 (0.022)	0.932 (0.018)	0.889 (0.098)	0.004 (0.002)	0.770 (0.193)	1.870 (0.583)	0.025 (0.010)	0.074 (0.012)
60	0.053 (0.018)	0.953 (0.015)	0.913 (0.085)	0.004 (0.002)	0.552 (0.141)	2.100 (0.604)	0.024 (0.009)	0.075 (0.011)
84	0.043 (0.016)	0.959 (0.014)	0.953 (0.082)	0.003 (0.002)	0.488 (0.125)	2.176 (0.608)	0.025 (0.008)	0.074 (0.010)
120	0.045 (0.019)	0.960 (0.016)	0.901 (0.071)	0.001 (0.001)	0.372 (0.102)	2.380 (0.01)	0.018 (0.01)	0.083 (0.011)

Notes: n: maturity in months, λ_n: feedback coefficient, ρ_n: risk premium AR(1) coefficient, b_n: policy MA coeffcient, σ^2: shock variances, standard errors in parentheses

Table 4: VECM results

All estimates are significantly positive, i.e. larger than two standard errors. The only exceptions are the innovation variances of the risk premia of the 7- and 10-year bonds. Obviously, ρ_n is rising with n, while λ_n is declining (except for the minor increase from seven to ten years). These results are in line with economic intuition: News on risk premia, which drive apart the yields of assets with different maturities, are the more persistent, the larger the difference between investment periods. This is because the fraction ρ of the risk premium that survives until the next period will be the higher the less a single period counts for the whole premium (for example only 1/120 for the ten-year bond); see Wolters and Weber (2009). Concerning the feedback coefficient, monetary policy reacts predominantly to the information contained in the spreads of shorter horizons, concentrating its forward-looking behaviour on near-future outlooks. Lastly, the MA coefficient b_n is bound to follow the long-run constraint.

The relative importance of the different innovations for the variation of the short- and long-term interest rates can be assessed by variance decompositions. This underlines the merits of the current approach to clearly separate the persistent and transitory components of monetary policy. The contemporaneous decompositions are calculated from the quadratic form $B_n Cov(e_t^{(n)} | t \in R_i) B_n'$, using the structural regime variances and the impact coefficients in B_n, arising from (7). Table 5 displays the variance contributions of the three shocks to both interest rates for all n in the second regime.

	long rate (n-month)			short rate (1-month)		
n	ε_1	ε_2	u	ε_1	ε_2	u
3	93.09	0.02	6.89	84.12	15.43	0.46
6	80.32	0.02	19.66	81.74	17.69	0.57
12	34.82	2.65	62.53	26.18	72.70	1.12
24	34.69	0.94	64.37	27.85	71.78	0.37
36	33.40	0.46	66.14	27.56	72.15	0.29
60	34.50	0.33	65.16	25.99	73.89	0.12
84	39.14	0.04	60.82	27.10	72.84	0.07
120	27.32	0.57	72.12	18.65	81.26	0.09

Table 5: Variance decompositions (2nd regime) in percent

The variance share explained by the risk premium shock is of minor importance for shorter-term interest rates, but rapidly rises until about two thirds for longer maturities. The share in the one-period rate variance is at most 1%. Concerning the policy shocks, it is logically the persistent one that dominates in the long rates. The same holds true for the one-period rate when paired with one of the shorter money market rates in one model. In contrast, for the cases of longer maturities, the largest part of the one-period rate is judged as transitory. This mirrors exactly the intended effect of driving a wedge into the one-to-one relationship implied by fully persistent shocks to the short end of the term structure. In particular, the rising importance of transitory policy shocks corresponds precisely to the rising evidence against the models with persistent policy innovations only in Table 2.

4 Conclusion

A recent strand of literature, including McCallum (2005), Kugler (1997) and Weber and Wolters (2009), analysed the term structure of interest rates in a model framework combining monetary policy and risk premium processes. Thereby, through arbitrage and rational expectations, persistent policy innovations to the short-term rate usually imply an immediate and undiminished reaction of the longer rates. In the underlying paper, we show that this hypothesis is empirically rejected for US interest rate data, making the implicit assumption of full persistence of policy shocks untenable.

To reconcile the theoretical model with empirical evidence, we introduce an additional policy shock, which is restricted to exert only transitory effects on the short-term rate. Since the variances of both the persistent and the transitory policy shock are freely estimated, their contributions are determined endogenously. Identification is achieved by exploiting the variance shift in interest rates after the Volcker disinflation period in the early 1980s. We derive the vector EC form from the structural equations and estimate the VECM for US term structure data.

Indeed, we find strong transitory policy effects in the term structure, driving a wedge between the responses of short and long interest rates. As could be expected, this wedge is the more important, the larger the difference in maturities. Furthermore, risk premia persistence is rising for longer-term rates. The estimates of the feedback coefficient in the stylised policy rule confirm that the Fed predominantly picks up information contained in spreads at the short end of the yield curve.

The current approach develops significant insights into the origins of the much-discussed empirical failure of the EHT. Beyond, it paves the way to a more sophisticated handling of the nature of monetary policy impulses to the economy. While we employed a highly stylised small-scale model, including policy shocks of different persistence in fully-fledged macro models appears both feasible and promising. In particular, influences of these shocks on key indicators like inflation and GDP growth could be investigated with respect to the type of policy impulse. We leave this for the agenda of future research.

References

[1] Ang, A., and M. Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, Journal of Monetary Economics, 60, 745–787.

[2] Campbell, J.Y., 1995, Some Lessons from the Yield Curve, Journal of Economic Perspectives, 9, 129-152.

[3] Campbell, J.Y. and R.J. Shiller, 1991, Yield spreads and interest rate movements: a bird's eye view, Review of Economic Studies, 58, 495-514.

[4] Caporale, B. and T. Caporale, 2008, Political risk and the expectations hypothesis, Economics Letters, 100, 178-180.

[5] Dai, Q., and K. Singleton, 2000, Specification Analysis of Affine Term Structure Models, Journal of Finance, 55, 1943-1978.

[6] Duffee, G.R., 2002, Term premia and interest rate forecasts in affine models, Journal of Finance, 57, 405–443.

[7] Engle, R.F. and C.W.J. Granger, 1987, Co-Integration and Error Correction: Representation, Estimation, and Testing, Econometrica, 55, 251-276.

[8] Engle, R.F., D.M. Lilien and R.P. Robins, 1987, Time Varying Risk Premia in the Term Structure: The Arch-M Model, Econometrica, 55, 391-407.

[9] Estrella, A., and G.A. Hardouvelis, 1991, The Term Structure as a Predictor of Real Economic Activity, Journal of Finance, 46, 555–76.

[10] Evans, M.D., and K.K. Lewis, 1994, Do Stationary Risk Premia Explain it all? Evidence from the Term Structure, Journal of Monetary Economics, 33, 285-318.

[11] Hamilton, J.D., 1988, Rational expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates, Journal of Economic Dynamics and Control, 12, 385-423.

[12] Hardouvelis, G.A., 1994, The Term Structure Spread and Future Changes in Long and Short Rates in the G7 Countries: Is there a Puzzle? Journal of Monetary Economics, 33, 255-283.

[13] Johansen, S., 1995, Likelihood-based Inference in Cointegrated Vector Autoregressive Models, Oxford: Oxford University Press.

[14] Johnson, M.H., 1988, Current Perspectives on Monetary Policy, Cato Journal, 8, 253-260.

[15] Juselius, K., 1999, Models and Relations in Economics and Econometrics, Journal of Economic Methodology, 6, 259–290.

[16] Kirchgässner G. and J. Wolters, 2007, Introduction to Modern Time Series Analysis, Berlin: Springer.

[17] Kugler, P., 1997, Central Bank Policy Reaction and the Expectations Hypothesis of the Term Structure, International Journal of Finance and Economics, 2, 217-224.

[18] Laurent, R.D., 1988, An Interest Rate-Based Indicator of Monetary Policy, Federal Reserve Bank of Chicago Economic Perspectives 12, 3–14.

[19] McCallum, B.T., 1983, On non-uniqueness in rational expectations models: An attempt at perspective, Journal of Monetary Economics, 11, 139-168.

[20] McCallum, B.T., 2005, Monetary Policy and the Term Structure of Interest Rates, Federal Reserve Bank of Richmond Economic Quarterly, 91, 1-21; first published in 1994 as NBER Working Paper 4938.

[21] Mankiw, N.G. and J.A. Miron, 1986, The changing behavior of the term structure of interest rates, Quarterly Journal of Economics, 101, 211-228.

[22] Mishkin, F.S., 1990, The Information in the Longer Maturity Term Structure about Future Inflation, Quarterly Journal of Economics, 105, 815–28.

[23] Poole, W., 1970, Optimal choice of monetary policy instruments in a simple stochastic macro model, Quarterly Journal of Economics, 84, 197-216.

[24] Rigobon, R., 2003, Identification through heteroscedasticity, Review of Economics and Statistics, 85, 777-792.

[25] Sentana, E., G. Fiorentini, 2001, Identification, estimation and testing of conditionally heteroskedastic factor models, Journal of Econometrics, 102, 143–164.

[26] Sola, M., J. Driffill, 1994, Testing the term structure of interest rates using a stationary vector autoregression with regime switching, Journal of Economics Dynamics and Control, 18, 601-628.

[27] Tzavalis, E. and M.R. Wickens, 1995, The persistence in volatility of the US term premium 1970-1986, Economics Letters, 49, 381-389.

[28] Tzavalis, E. and M.R. Wickens, 1998, A Re-Examination of the Rational Expectations Hypothesis of the Term Structure: Reconciling the Evidence from Long-Run and Short-Run Tests, International Journal of Finance and Economics, 3, 229-239.

[29] Weber, E., J. Wolters, 2009, The US Term Structure and Central Bank Policy, Universität Regensburg, Economics Discussion Paper 436.

5 Appendix

In the following we shall derive the VECM representation of our preferred model consisting of the modified EHT of equation (1) and the generalised policy rule in equation (6). Assuming that expectations are formed rationally, the model can be solved using the method of undetermined coefficients (see McCallum 1983). The minimal state variable solution is then given by

$$r_t = \phi_1^{(n)} r_{t-1} + \phi_2^{(n)} v_t^{(n)} + \phi_3^{(n)} \varepsilon_{1t}^{(n)} + \phi_4^{(n)} q_t^{(n)} + \phi_5^{(n)} q_{t-1}^{(n)} \; . \tag{10}$$

Since r_t behaves like an I(1)-process, $\phi_1^{(n)}$ has to take on the value 1. For the other undetermined coefficients one gets

$$\phi_2^{(n)} = \frac{\lambda_n n}{n - \lambda_n \rho_n S(\rho_n)} \; , \; \phi_3^{(n)} = 1 \; , \; \phi_4^{(n)} = \frac{n - \lambda_n b_n S(a_n)}{n - \lambda_n a_n S(a_n)} \; , \; \phi_5^{(n)} = -b_n$$

with $S(x) := \sum_{j=1}^{n-1} (n-j) x^{j-1}$.

Substituting in (10), we obtain

$$\Delta r_t = \frac{\lambda_n n}{n - \lambda_n \rho_n S(\rho_n)} v_t^{(n)} + \varepsilon_{1t}^{(n)} + \frac{n - \lambda_n b_n S(a_n)}{n - \lambda_n a_n S(a_n)} q_t^{(n)} - b_n q_{t-1}^{(n)} \tag{11}$$

and then using equation (6),

$$R_t^{(n)} - r_t = \frac{n}{n - \lambda_n \rho_n S(\rho_n)} v_t^{(n)} + \frac{(a_n - b_n) S(a_n)}{n - \lambda_n a_n S(a_n)} q_t^{(n)} \; . \tag{12}$$

From (11) and (12) we derive the following short-run effects of the three shocks:

$$\frac{\partial r_t}{\partial u_t^{(n)}} = \frac{\lambda_n n}{n - \lambda_n \rho_n S(\rho_n)} \; , \; \frac{\partial r_t}{\partial \varepsilon_{1t}^{(n)}} = 1 \; , \; \frac{\partial r_t}{\partial \varepsilon_{2t}^{(n)}} = \frac{n - \lambda_n b_n S(a_n)}{n - \lambda_n a_n S(a_n)}$$

$$\frac{\partial R_t^{(n)}}{\partial u_t^{(n)}} = \frac{(1+\lambda_n) n}{n - \lambda_n \rho_n S(\rho_n)} \; , \; \frac{\partial R_t^{(n)}}{\partial \varepsilon_{1t}^{(n)}} = 1 \; , \; \frac{\partial R_t^{(n)}}{\partial \varepsilon_{2t}^{(n)}} = \frac{n - \lambda_n b_n S(a_n) + (a_n - b_n) S(a_n)}{n - \lambda_n a_n S(a_n)}$$

The ARMA representation of (12) ist given as

$$R_t^{(n)} - r_t = \rho_n (R_{t-1}^{(n)} - r_{t-1}) + \frac{n}{n - \lambda_n \rho_n S(\rho_n)} u_t^{(n)} + \frac{(a_n - b_n) S(a_n)}{n - \lambda_n \rho_n S(\rho_n)} (q_t^{(n)} - \rho_n q_{t-1}^{(n)}) \; . \tag{13}$$

Equating (12) and (13), solving for $v_t^{(n)}$ and substituting in (11) leads to an error-correction equation for the short-term interest rate with MA(1)-residuals:

$$\begin{aligned}\Delta r_t &= \lambda_n\rho_n(1-a_n)(R^{(n)}_{t-1}-r_{t-1})+a_n(1-\lambda_n\rho_n)\Delta r_{t-1}+a_n\lambda_n\rho_n\Delta R^{(n)}_{t-1}\\ &+\frac{\lambda_n n}{n-\lambda_n\rho_n S(\rho_n)}(u^{(n)}_t-a_n u^{(n)}_{t-1})+\varepsilon^{(n)}_{1t}-a_n\varepsilon^{(n)}_{1t-1}\\ &+\frac{n-\lambda_n b_n S(a_n)}{n-\lambda_n a_n S(a_n)}\varepsilon^{(n)}_{2t}+\frac{\lambda_n S(a_n)(a_n b_n-\rho_n(a_n-b_n))-b_n n}{n-\lambda_n a_n S(a_n)}\varepsilon^{(n)}_{2t-1}\end{aligned} \tag{14}$$

Rearranging (13) and substituting (14) gives the error-correction equation of the long-term interest rate

$$\begin{aligned}\Delta R^{(n)}_t &= (\lambda_n\rho_n+\rho_n-1)(1-a_n)(R^{(n)}_{t-1}-r_{t-1})+a_n(1-\rho-\lambda_n\rho_n)\Delta r_{t-1}+a_n\rho_n(1+\lambda_n)\Delta R^{(n)}_{t-1}\\ &+\frac{(1+\lambda_n)n}{n-\lambda_n\rho_n S(\rho_n)}(u^{(n)}_t-a_n u^{(n)}_{t-1})+\varepsilon^{(n)}_{1t}-a_n\varepsilon^{(n)}_{1t-1}\\ &+\frac{n+S(a_n)(a_n-b_n-\lambda_n b_n)}{n-\lambda_n a_n S(a_n)}\varepsilon^{(n)}_{2t}+\frac{S(a_n)(\lambda_n a_n b_n-\rho_n(a_n-b_n)(1+\lambda_n))-b_n n}{n-\lambda_n a_n S(a_n)}\varepsilon^{(n)}_{2t-1}\,.\end{aligned} \tag{15}$$

For the special case $a_n = 0$ we get from (14) and (15):

$$\begin{aligned}\Delta r_t &= \lambda_n\rho_n(R^{(n)}_{t-1}-r_{t-1})+\frac{\lambda_n n}{n-\lambda_n\rho_n S(\rho_n)}u^{(n)}_t+\varepsilon^{(n)}_{1t}\\ &+(1-\lambda_n b_n\frac{n-1}{n})\varepsilon^{(n)}_{2t}+(\lambda_n\rho_n\frac{n-1}{n}-1)b_n\varepsilon^{(n)}_{2t-1}\end{aligned} \tag{16}$$

$$\begin{aligned}\Delta R^{(n)}_t &= (\lambda_n\rho_n+\rho_n-1)(R^{(n)}_{t-1}-r_{t-1})+\frac{(1+\lambda_n)n}{n-\lambda_n\rho_n S(\rho_n)}u^{(n)}_t+\varepsilon^{(n)}_{1t}\\ &+(1-\frac{n-1}{n}(1+\lambda_n)b_n)\varepsilon^{(n)}_{2t}+(\frac{n-1}{n}(1+\lambda_n)\rho_n-1)b_n\varepsilon^{(n)}_{2t-1}\,.\end{aligned} \tag{17}$$