



Intelligible Factors for the Yield Curve

Yvan Lengwiler and Carlos Lenz

The views expressed in this paper are those of the author(s) and do not necessarily represent those of the Swiss National Bank. Working Papers describe research in progress. Their aim is to elicit comments and to further debate.

ISSN 1660-7716 (printed version)
ISSN 1660-7724 (online version)

© 2008 by Swiss National Bank, Börsenstrasse 15, P.O. Box, CH-8022 Zurich

Intelligible Factors for the Yield Curve*

Yvan Lengwiler[†] Carlos Lenz[‡]

November 29, 2007

Abstract

We construct a factor model of the yield curve and specify time series processes for these factors, so that the innovations are mutually orthogonal. At the same time, the factors are constructed in such a way that they assume clear, intuitive interpretations. The resulting “intelligible factors” should prove useful for investment professionals to discuss expectations about yield curves and the implied dynamics. Moreover, they allow us to distinguish announced changes of the monetary policy stance versus monetary policy surprises, which are actually rare. We identify two such events, namely September 11, 2001, and the Fed reaction to the recent sub-prime crisis.

JEL classification: E43.

Keywords: yield curve, factor models, structural vector autoregression, monetary policy.

© Copyright by the authors.

*We thank Paul Söderlind and the participants of Bundesbank-OeNB-SNB workshop 2007 as well as an anonymous referee for useful comments. The views expressed in this paper do not necessarily reflect those of the Swiss National Bank.

[†]Department of Economics (WWZ), University of Basel, Petersgraben 51, CH-4003 Basel, Switzerland, email: yvan.lengwiler@unibas.ch

[‡]Research Unit, Swiss National Bank, Börsenstrasse 15, P.O. Box, CH-8022 Zürich, Switzerland, email: carlos.lenz@snb.ch

1. Introduction

For people who are required to form expectations about complex objects, such as the term structure of interest rates, it is useful to reduce the dimensionality of the problem with the help of a few factors that describe the object. Ideally, these factors should have two properties. Firstly, they should have an *interpretation* that is easy to understand. For instance, when we talk about the yield curve, the short rate is an interpretable factor, and the slope or the long rate, too, are interpretable factors. We can discuss our expectations about these things and it is clear what it means. Secondly, the factors should be driven by processes that have innovations that are *mutually orthogonal*. It is difficult to form expectations about the long rate and the slope separately if the long rate and the slope are driven by shocks that are not independent from each other. In this paper, we say that factors that fulfill these two requirements of being interpretable and having mutually and serially orthogonal innovations, are *intelligible*.

The literature on the term structure of interest rates has produced a plethora of factor models, which can be divided into three lines. The first consists of models that start with a specification of some stochastic process for the short rate. They then derive the dynamics of the term structure by imposing arbitrage-freeness. Early representatives in this line of research are Pye (1966), Vasicek (1977), and Cox, Ingersoll, and Ross (1981). These models have been extended to multi-factor models where yields depend on the factors in an affine fashion. Duffie and Kan (1996) and Dai and Singleton (2000) are two prominent examples in this domain. In these models, the factors are typically some abstract entities with no clear economic interpretation.

The second line consists of general equilibrium models that explain the term structure from first principles. The classic reference here is Cox, Ingersoll, and Ross (1985). They explicitly derive an equilibrium term structure process as a function of the preferences of a representative utility maximizer and an assumed process of k factors that describe the production possibilities of the economy. These k factors have macroeconomic interpretations by construction. This model, therefore, is an early representative of the macro-finance literature that explores the relation between asset prices in general — or in this case, the term structure in particular — with observable or latent macroeconomic variables. The modern version of this literature (e.g. Ang and Piazzesi, 2003, Rudebusch and Wu, 2004, Mönch, 2005) is much less explicit about the general equilibrium underpinnings of the model. It takes up the idea that the short rate depends on a set of macroeconomic variables and imposes arbitrage-freeness to derive the implications for the relation between the term structure and these macroeconomic factors. By using exogenous variables to estimate the model, they achieve a better fit of the yield curve compared to the models with unspecified factors.

The third line of the literature consists of purely descriptive, empirical factor models. These models describe the term structure with the help of a few factors in order to facilitate communication about it. This literature is motivated more by the needs of the practitioner than by the interests of the economic theorist. A classic early refer-

ence to this type of literature is Litterman and Scheinkman (1991), who apply principal component analysis to yield curve data. This analysis generates orthogonal factors by construction. Typically, three factors suffice to describe the yield curve, and the corresponding loadings suggest an interpretation as ‘level,’ ‘slope,’ and ‘curvature.’ Another early contribution in this domain is Nelson and Siegel (1987). These contributions are completely static in the sense that they model the yield curve at a particular point in time. They do not, however, contain information about the dynamics of the yield curve and can therefore not be used for forecasting. Their main advantage over the theory-based models is their better fit. Diebold and Li (2006) have combined the Nelson-Siegel model with an autoregressive or vector autoregressive (VAR) specification for the factor process. They show that this combination of a set of easily interpretable factor loadings together with a simple stochastic process yields better forecasting performance than the dynamic theory-based models. The drawback is, of course, that it is not guaranteed that the model is arbitrage-free. Moreover, the factor innovations in the Diebold-Li model are not independent of each other. In other words, in the Nelson-Siegel model, one cannot discuss level-innovations independently from slope-innovations because these innovations are statistically not orthogonal. This jeopardizes interpretability. One can still use these models for forecasting, but the meaning of shocks to the individual factors is unclear.

In this paper, we address this issue. We construct a factor model in the tradition of the third line of research, but impose orthogonality of the factor innovations in addition to a clear interpretation of the loadings, i.e. we restrict our factors to be intelligible. More concretely, our loadings, which can be identified as ‘long’, ‘short,’ and ‘curvature’ factors, follow a VAR process with mutually independent and serially uncorrelated innovations. In other words, we construct the factors in such a way that the innovations to our short factor, for instance, are uncorrelated with innovations to the other two factors.

In addition to the imposed intelligibility of the factors, our estimation reveals a dynamic structure which suggests a macroeconomic interpretation. We find that the curvature factor is a leading indicator of the short factor, while the long factor largely lives a life of its own. We will argue that this dynamic structure suggests that the curvature factor captures the intended (and communicated) medium term monetary policy stance, and the short factor captures surprise policy actions. This relates our model to the macro-finance term structure literature. We find that the curvature factor explains a much greater share of term structure movements than the other two factors. This suggests that most monetary policy actions are expected (maybe because they are well communicated by the Fed) before they are actually executed. We do, however, identify two significant exceptions to this rule.

2. The model

Our basic approach consists of two steps: First, we formulate an interpretable factor model for the yield curve at one point in time. Second, we look at the common evolu-

tion of the factors over time and impose a dynamic structure which makes the factors intelligible.

2.1. A parsimonious model

We will use two related factor models for the yield curve. Let the raw factors $\theta = [\theta_1 \dots \theta_F]'$ be an F -dimensional column vector. Define their loadings $y(m) = [y_1(m) \dots y_F(m)]$ as an F -dimensional function,

$$y_i(m) = \begin{cases} -\frac{1-\alpha_i^m}{m}, & \text{if } m > 0, \\ \ln \alpha_i, & \text{if } m = 0, \end{cases} \quad (1)$$

with $0 < \alpha_i < 1$ for all $i \in \{2, \dots, F\}$, and $y_1(m) = 1$ for all m .¹ The loadings y are continuous monotonic concave functions of the maturity with negative values that converge to zero as $m \rightarrow \infty$. The yield curve at time t is described by

$$r_t = y\theta_t + \varepsilon_t = \hat{r}_t + \varepsilon_t, \quad (2)$$

where ε_t is the idiosyncratic and \hat{r}_t the systematic component. We will only consider the model with $F = 3$,

$$\hat{r}_t = \theta_{1,t} + y_2\theta_{2,t} + y_3\theta_{3,t}. \quad (3)$$

Given yield curve data of a day, and given α , the factors of that day, θ_t , can be estimated by OLS. The best choice of α can be estimated with a non-linear regression. Once we know the factors for each point in time, $\theta = [\theta_1 \dots \theta_T]$, and building on the idea of Diebold and Li (2006) we can estimate a VAR model for θ .² The VAR model in the factors represents, together with the loadings y (parametrized by α), a complete description of the yield curve dynamics. Therefore, it allows, among other things, to assess the dynamic reaction of the yield curve to factor innovations, which is potentially useful if these innovations can be given a structural interpretation.

However, θ are not intelligible factors because they are difficult to interpret. Moreover, there is nothing that guarantees that the innovations in θ are mutually orthogonal. Consider therefore a transformation of θ ,

$$\phi = B\theta, \quad (4)$$

¹The functional form of these loadings is just one example of loadings that can be used for what we try to accomplish. We will later discuss some of the properties that these loadings must have, and especially, why the traditional Nelson and Siegel (1987) specification of the loadings, or the various modifications of it (Svensson, 1994, Bliss, 1997, De Pooter, 2007), are not suitable for our purpose.

²Diebold and Li (2006) focus primarily on univariate AR-models for each factor separately as they observe little cross-factor interaction.

and define the corresponding loadings

$$k(m) = y(m)B^{-1}, \quad (5)$$

assuming that B is invertible. Then clearly $\hat{r} = k\phi$. We want to choose B such that the transformed factors ϕ are interpretable, and the innovations to these factors are orthogonal.

2.2. Making the factors intelligible

Consider first the long and the short ends of the yield curve. Because the two ends of the yield curve are easy to interpret, we want the first two transformed factors to represent just this, i.e. ϕ_1 should represent the *long rate* and ϕ_2 should represent the *short rate*. This requires that the transformed factor loadings satisfy

$$\lim_{m \rightarrow \infty} k(m) = [1 \ 0 \ 0], \quad (6)$$

$$\lim_{m \rightarrow 0} k(m) = [0 \ 1 \ 0]. \quad (7)$$

This means that the first factor, ϕ_1 , is the only one that is affected by changes of the long rate (changes of the yield with infinite maturity), and the second factor, ϕ_2 , is the only one that is affected by changes of the short rate (changes of the yield with imminent maturity). These interpretability constraints guarantee that the interpretation of ϕ_1 as the long and ϕ_2 as the short rate are sensible.

If all $\alpha_i < 1$ we have³

$$\lim_{m \rightarrow \infty} \hat{r}(m) = \theta_1, \quad (8)$$

$$\lim_{m \rightarrow 0} \hat{r}(m) = \theta_1 + \ln \alpha_2 \theta_2 + \ln \alpha_3 \theta_3. \quad (9)$$

Given these facts, (6) and (7) restrict some components of B . Using the definition of the transformed loadings (5), B has to satisfy

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & \ln \alpha_2 & \ln \alpha_3 \end{bmatrix}. \quad (10)$$

This restricts two of the three rows of B ,

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & \ln \alpha_2 & \ln \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{bmatrix}. \quad (11)$$

The coefficients β determine the linear combination of the raw factors θ that generate

³ $\lim_{m \rightarrow \infty} \hat{r}(m)$ is not defined if some $\alpha_i > 1$.

the third transformed factor ϕ_3 , which we construct in such a way that it can be interpreted as *curvature*. The restrictions (6) and (7) ensure the loading of ϕ_3 to vanish as m converges to either zero or infinity. We furthermore normalize β so that the maximum loading on ϕ_3 (i.e. $\max_m k_3(m)$) is unity. As a consequence, the loading of the third factor begins at zero, reaches unity at some intermediate maturity, and then converges asymptotically back to zero.

B contains five coefficients. Of these five coefficients, one is used for the normalization of the loading of the third factor. Three coefficients are used to make sure that the innovations to the transformed factors are orthogonal. This leaves us with one free coefficient, which is used to maximize the fit of the model.

In order to define innovations to the factors, we need to specify a stochastic process. We choose a VAR specification which we fit to the estimated factors θ ,

$$\theta_t = A_0 + A_1\theta_{t-1} + A_2\theta_{t-2} + \dots + A_p\theta_{t-p} + v_t. \quad (12)$$

$\Omega = E[v_t v_t']$ is the covariance matrix of the residuals. There is nothing that guarantees that the residuals are independent, i.e. that Ω is diagonal. In fact, our estimates reveal these residuals to be far from orthogonal. We transform this VAR into a structural VAR model by using our matrix B ,

$$B\theta_t = BA_0 + BA_1B^{-1}B\theta_{t-1} + BA_2B^{-1}B\theta_{t-2} + \dots + BA_pB^{-1}B\theta_{t-p} + Bv_t. \quad (13)$$

Noting that $B\theta = \phi$ and defining $D_0 = BA_0$, $D_i = BA_iB^{-1}$ for $i = 1, \dots, p$, and $u_t = Bv_t$, reveals that (13) is a VAR in the transformed factors ϕ ,

$$\phi_t = D_0 + D_1\phi_{t-1} + D_2\phi_{t-2} + \dots + D_p\phi_{t-p} + u_t. \quad (14)$$

The innovations u_t in this VAR model are made mutually orthogonal through the choice of B . Therefore, they can be given a structural interpretation as innovations to the long, short, and curvature factors. The covariance matrix of u_t equals $B\Omega B'$ and has the following structure

$$B\Omega B' = \begin{bmatrix} \Omega_{1,1} & z\Omega_1 & \beta\Omega_1 \\ z\Omega_1 & z\Omega z' & \beta\Omega z' \\ \beta\Omega_1 & \beta\Omega z' & \beta\Omega\beta' \end{bmatrix}, \quad (15)$$

where $z = [1 \ \ln \alpha_2 \ \ln \alpha_3]$ and Ω_i denotes the i 'th column of Ω . The required orthogonality of the structural innovations imposes that $B\Omega B'$ be diagonal which places three restrictions on α and β .

Finding the right β is simple because β does not affect Ω . It is simply a basis for the null-space of $[\Omega_1 \ \Omega z']$. These two restrictions together with the normalization mentioned before fix β . We determine α_2 in such a way that the covariance matrix becomes orthogonal ($z\Omega_1$ vanishes). This is not so straightforward, though, because α_2

affects θ and thus also alters Ω . The right α_2 can be found iteratively, however. The remaining coefficient α_3 can be chosen to maximize the fit of the model.

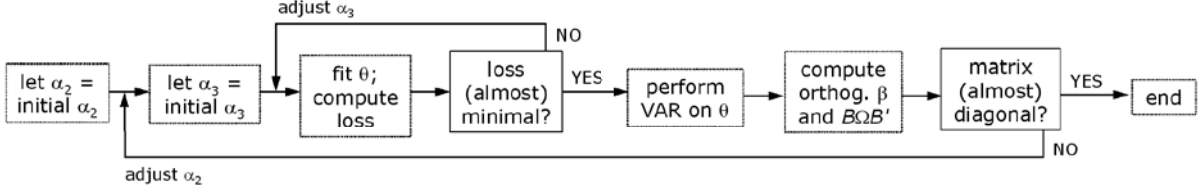


Figure 1: Structure of the nested optimization.

We are therefore confronted with a nested optimization problem, as depicted in Figure 1. The “inner optimization” selects α_3 so that the fit of the model is maximized. We choose α_3 numerically to minimize the objective function

$$\text{loss} = \frac{1}{T} \sum_{t=1}^T \frac{\varepsilon'_t \varepsilon_t}{M_t - 3}, \quad (16)$$

that is, our objective is the average adjusted mean squared error of the yield curve model (2) fitted to the data of day t . M_t is the number of maturities we observe in period t . The loss function also depends on α_2 , but this is given at that point. The “outer optimization” selects α_2 so that the covariance-matrix of u_t , eq. (15), is diagonal. In particular, we use a root search algorithm that selects α_2 so that the correlation of the innovations to ϕ_1 and ϕ_2 vanishes. The other two correlations are taken care of already with the selection of β .⁴

2.3. Revisiting Nelson-Siegel

Before presenting our numerical results it may be in order to make a quick digression. The Nelson and Siegel (1987) model has been applied quite successfully by practitioners and academics because it provides enough flexibility to describe the yield curve rather well, but is unable to produce too crazy shapes which never occur empirically. In our notation, the Nelson-Siegel model amounts to an alternative definition of θ and y . We add tildes on top of the variables whenever we refer to the Nelson-Siegel model. The

⁴Note that the procedure does not yield a unique solution in a trivial fashion. If the outer optimization converges to some number α_2 and the inner converges to some number α_3 , then switching the two numbers is also a solution.

Nelson-Siegel model has $F = 3$ with loadings defined as

$$\tilde{y}_1(m) = 1, \quad (17)$$

$$\tilde{y}_2(m) = -\frac{1 - \exp(-m/\tau)}{m/\tau}, \quad (18)$$

$$\tilde{y}_3(m) = \frac{1 - \exp(-m/\tau)}{m/\tau} - \exp(-m/\tau), \quad (19)$$

where $\tau > 0$ is a parameter. The Nelson-Siegel yield curve is then

$$r = \tilde{y}\tilde{\theta} + \tilde{\varepsilon}. \quad (20)$$

In the Nelson-Siegel model, the long and the short rates are defined, respectively, as

$$\lim_{m \rightarrow \infty} r(m) = \tilde{\theta}_1, \quad (21)$$

$$\lim_{m \rightarrow 0} r(m) = \tilde{\theta}_1 - \tilde{\theta}_2. \quad (22)$$

We can estimate this model and also determine the time series properties of the $\tilde{\theta}$ -factors. However, it is not possible to specify these processes in such a way that the innovations to these factors are mutually orthogonal (just as it is not possible to do this with the θ -processes). In fact, empirically, all three factors have innovations that are statistically dependent on the other two factors. As before, we search for a transformation of the factors that would make the innovations orthogonal, but that keeps their interpretability intact at the same time.

(21) and (22) imply the following interpretability restrictions on the transformation matrix,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tilde{B} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}. \quad (23)$$

As a result, \tilde{B} has the following structure,

$$\tilde{B} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ \tilde{\beta}_1 & \tilde{\beta}_2 & \tilde{\beta}_3 \end{bmatrix}. \quad (24)$$

The covariance matrix of the structural VAR has the following form,

$$\tilde{B}\tilde{\Omega}\tilde{B}' = \begin{bmatrix} \tilde{\Omega}_{1,1} & \tilde{\Omega}_{1,1} - \tilde{\Omega}_{1,2} & \tilde{\beta}\tilde{\Omega}_1 \\ \tilde{\Omega}_{1,1} - \tilde{\Omega}_{2,1} & \tilde{\Omega}_{1,1} + \tilde{\Omega}_{2,2} & \tilde{\beta}(\tilde{\Omega}_1 - \tilde{\Omega}_2) \\ \tilde{\beta}\tilde{\Omega}_1 & \tilde{\beta}(\tilde{\Omega}_1 - \tilde{\Omega}_2) & \tilde{\beta}\tilde{\Omega}\tilde{\beta}' \end{bmatrix}. \quad (25)$$

We see from this result that we can select $\tilde{\beta}$ in such a way that the alternative curvature

factor, $\tilde{\phi}_3 = \tilde{\beta}\tilde{\theta}$, has innovations that are orthogonal on the innovations of the other two factors. This provides some enhancement of the Nelson-Siegel model. However, it is not possible within the Nelson-Siegel specification to formulate long and short factors that have orthogonal innovations. $\tilde{B}\tilde{\Omega}\tilde{B}'$ cannot be made diagonal because the model has no free coefficient on all off-diagonal components in (25). It is therefore not possible to transform the Nelson-Siegel factors in such a way to make them intelligible.

The reason why the Nelson-Siegel specification of the loadings is not suitable for making the factors intelligible lies in their specific functional form, (18) and (19). In particular, the factors are constructed in such a way that $\lim_{m \rightarrow 0} \tilde{y}_2(m) = -1$ and $\lim_{m \rightarrow 0} \tilde{y}_3(m) = 0$, and $\lim_{m \rightarrow \infty} \tilde{y}_2(m) = \lim_{m \rightarrow \infty} \tilde{y}_3(m) = 0$. All these limits are independent of the only coefficient of the Nelson-Siegel model τ . The interpretability constraints (23), however, impose restrictions on \tilde{B} only with respect to the loadings at zero and at infinite maturity. This is why τ drops out of the rotation matrix \tilde{B} (24), and can thus not be used to make the innovations into the factors mutually orthogonal. The same would apply to the four-factor extension proposed by Svensson (1994) (or to similar modifications as in De Pooter, 2007), which features two different curvature factors, because the loadings of the Nelson-Siegel-Svensson model also converge to values which are independent of the coefficients of the model as m assumes an extreme value (zero or infinity).

3. Results

3.1. Data

We use the daily fixed maturity yields provided by the Federal Reserve on its internet site.⁵ We consider daily observations from the beginning of January 1990 to the end of October 2007. This is 4463 days worth of data. The data contain information for eleven maturities, 1 month, 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, 7 years, 10 years, 20 years, and 30 years. Most maturities are available for all days in our sample, except the 1 month rate which is available only from July 31, 2001 on, the 20 years rate which is available only since October 1, 1993, and the 30 year rate, which is not available from February 19, 2002 to February 8, 2006. Figure 2 depicts a longitudinal plot of all the data, as well as a few yield curves for some sample dates.

3.2. Estimation

There is no theoretical guarantee that the nested optimization works, but with our data, it appears to work well.⁶ We do find a VAR-specification that produces serially

⁵See section "Treasury constant maturities, Nominal" on <http://www.federalreserve.gov/releases/h15/data.htm>, also available in a more convenient form from the Treasury, <http://www.treas.gov/offices/domestic-finance/debt-management/interest-rate/yield.shtml>.

⁶All estimations are performed with MATLAB.

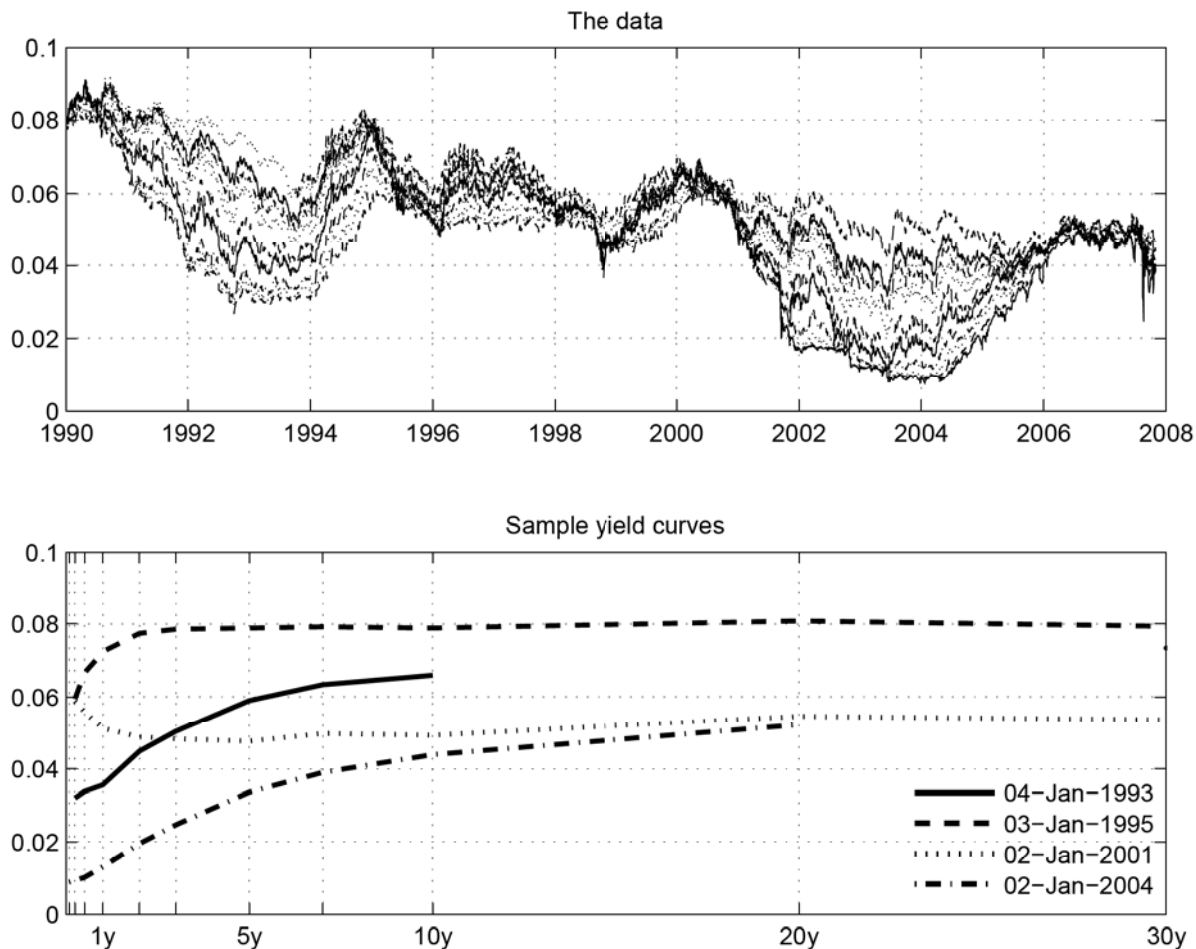


Figure 2: A first view of the data.

uncorrelated innovations, and we also find coefficients that make the innovations to the ϕ -processes mutually orthogonal and B invertible. The precise results depend on the VAR-specification, of course. We find that a VAR with 30 lags is sufficient to get rid of autocorrelation, see Table 1. This rather large number of lags is due mainly to the autocorrelation of ε_2 , the innovations to the short rate. 30 lags (somewhat more than one month worth of data) has a certain logic given the frequency of FOMC decisions.

The nested optimization yields the following coefficients,

$$\begin{aligned}
 \alpha_2 &= 0.1133, & \beta_1 &= 0.2674, \\
 \alpha_3 &= 0.6798, & \beta_2 &= -0.4343, \\
 & & \beta_3 &= -0.2584.
 \end{aligned}$$

Table 1: Box-Ljung Q-statistics (p values in parentheses).

lag	u_1		u_2		u_3	
1	0.0002	(0.99)	0.0051	(0.94)	0.0125	(0.91)
2	0.0003	(1.00)	0.0381	(0.98)	0.0127	(0.99)
3	0.0058	(1.00)	0.0423	(1.00)	0.0144	(1.00)
4	0.0238	(1.00)	0.0515	(1.00)	0.0151	(1.00)
5	0.0362	(1.00)	0.0708	(1.00)	0.0224	(1.00)
10	0.0494	(1.00)	0.1124	(1.00)	0.0942	(1.00)
15	0.2998	(1.00)	0.4310	(1.00)	0.2608	(1.00)
20	0.7488	(1.00)	1.8483	(1.00)	0.5103	(1.00)
30	1.2628	(1.00)	8.8520	(1.00)	1.6664	(1.00)
40	7.7539	(1.00)	29.7599	(0.88)	7.7989	(1.00)
50	19.3904	(1.00)	50.8359	(0.44)	24.5693	(1.00)
60	34.1828	(1.00)	66.7644	(0.26)	32.3741	(1.00)
70	50.6608	(0.96)	81.2480	(0.17)	54.7025	(0.91)
80	59.1151	(0.96)	99.8393	(0.07)	65.0302	(0.89)
100	81.0455	(0.92)	112.122	(0.19)	78.6296	(0.94)
120	100.504	(0.90)	133.991	(0.18)	98.4937	(0.92)

The minimized loss is $1.145 \cdot 10^{-6}$. This corresponds to an adjusted root mean squared error of 10 basis points (bp) on average over all days. The largest average error on any day is 35 bp (see Figure 3).

The mutual correlation coefficients of the innovations into the factor processes, u (14), essentially vanish. The absolute correlations of u_3 with either u_1 or u_2 are less than $2 \cdot 10^{-15}$; the correlation coefficient of u_1 and u_2 is $1.6 \cdot 10^{-8}$.

In contrast, if we perform a VAR on estimated Nelson-Siegel factors, the innovations into these processes are clearly not orthogonal. The three factors of this model are level, slope, and curvature. The innovations into the level and the slope factors correlate with a coefficient of 0.67, the correlation between level innovation and curvature innovation is 0.20, and between slope innovation and curvature innovation it is 0.39. It is therefore not so clear what we mean, for instance, by a “shock to the level” within the Nelson-Siegel model, since such shocks typically occur together with a “shock to the slope.”

3.3. The factors and their loadings

The right column of Figure 4 depicts the loadings k of the three transformed factors ϕ . These loading depend on the estimated α - and β -vectors. The implied loadings allow for rather clear interpretations. The loading of ϕ_2 is (by construction) unity at maturity zero. It declines very rapidly for positive maturities. It actually becomes slightly negative for times to maturity beyond 2.5 years, but is very close to zero for longer maturities.

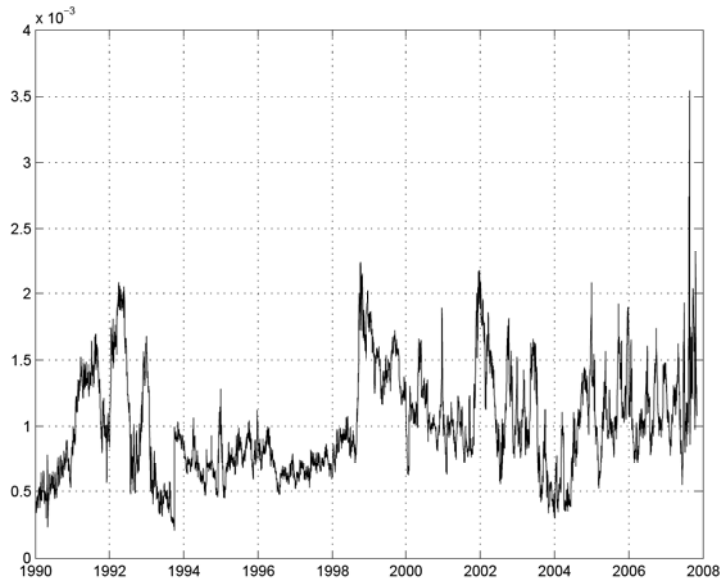


Figure 3: Adjusted root mean squared error for each day.

ϕ_2 thus essentially represents the money market part of the yield curve. The loading of the curvature factor, ϕ_3 , is also quite clear: it starts at zero (by construction), peaks a little bit before 2 years where it reaches unity, and then declines slowly as the time to maturity rises further. The loading at 9 years is slightly below 0.5, at 23 years it is 0.2. Finally, the loading of the long factor, ϕ_1 , also starts at zero, but then rapidly increases with the time to maturity. It reaches 0.5 at about 1 year and 0.8 around 3 years. It then slowly converges to one as we increase the time to maturity. A ϕ_1 -innovation thus manifests itself not only at the very long end of the yield curve, but is quite noticeable also at intermediate horizons.

The left column of Figure 4 depicts the estimated factors ϕ for each day in our sample. The development of ϕ_2 essentially mirrors the FOMC decisions. It strongly correlates with the Federal Funds Rate (though not perfectly so). There is a striking similarity between ϕ_2 and ϕ_3 . Remember that the innovations to these processes are orthogonal by construction, so the obvious correlation between the factors is due to the fact that they are coupled to each other through the dynamics implied by the VAR. Finally, ϕ_1 is an interesting measure of the long term interest rate. According to this measure, the long rate has decreased from roughly 9% to about 4.5%. During the time span covered by our data, core inflation (measured as the annual increase of the consumer price index less food and energy) has decreased from about 5% to 2%. This fact, together with our estimate of the ϕ_1 factor, implies that the long horizon *real* interest rate (measured with the Fisher equation, assuming that expected inflation has been identical to realized inflation) has decreased from 4.0% to about 2.5% during that period.

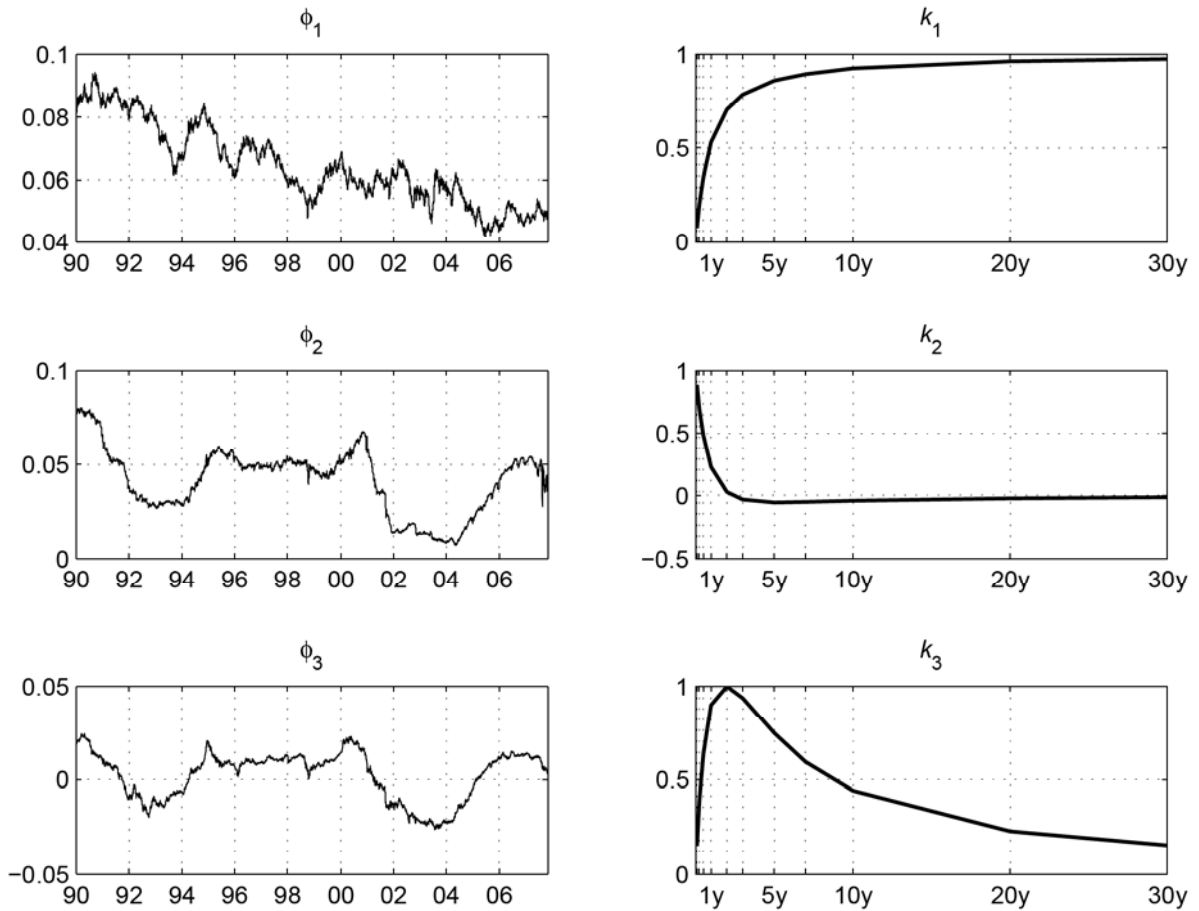


Figure 4: Estimated rotated factors (ϕ) and their loadings (k).

3.4. Impulse-response analysis in factor space

The impulse-response analysis reveals that innovations die out only quite slowly, see Figure 5. The long factor (ϕ_1) seems so be moved more or less only by own innovations; see the left column of Figure 5. It takes about 500 days for an innovation to ϕ_1 to reduce to half of its size, so innovations to the long factor are very persistent. Innovations to the long factor also drag along the short factor (ϕ_2) and, to a smaller extent, the curvature factor (ϕ_3); see the top row of Figure 5.

Innovations to the short factor essentially affect the short factor only. This kind of innovation takes about 60 days to reduce to half its size. There is very little spillover to the other two factors, except over very short horizons. On the other hand, the short factor is strongly influenced by innovations to the long factor, and even more so by innovations to the curvature factor.

The curvature factor has a surprising dynamics. A unit innovation cumulatively

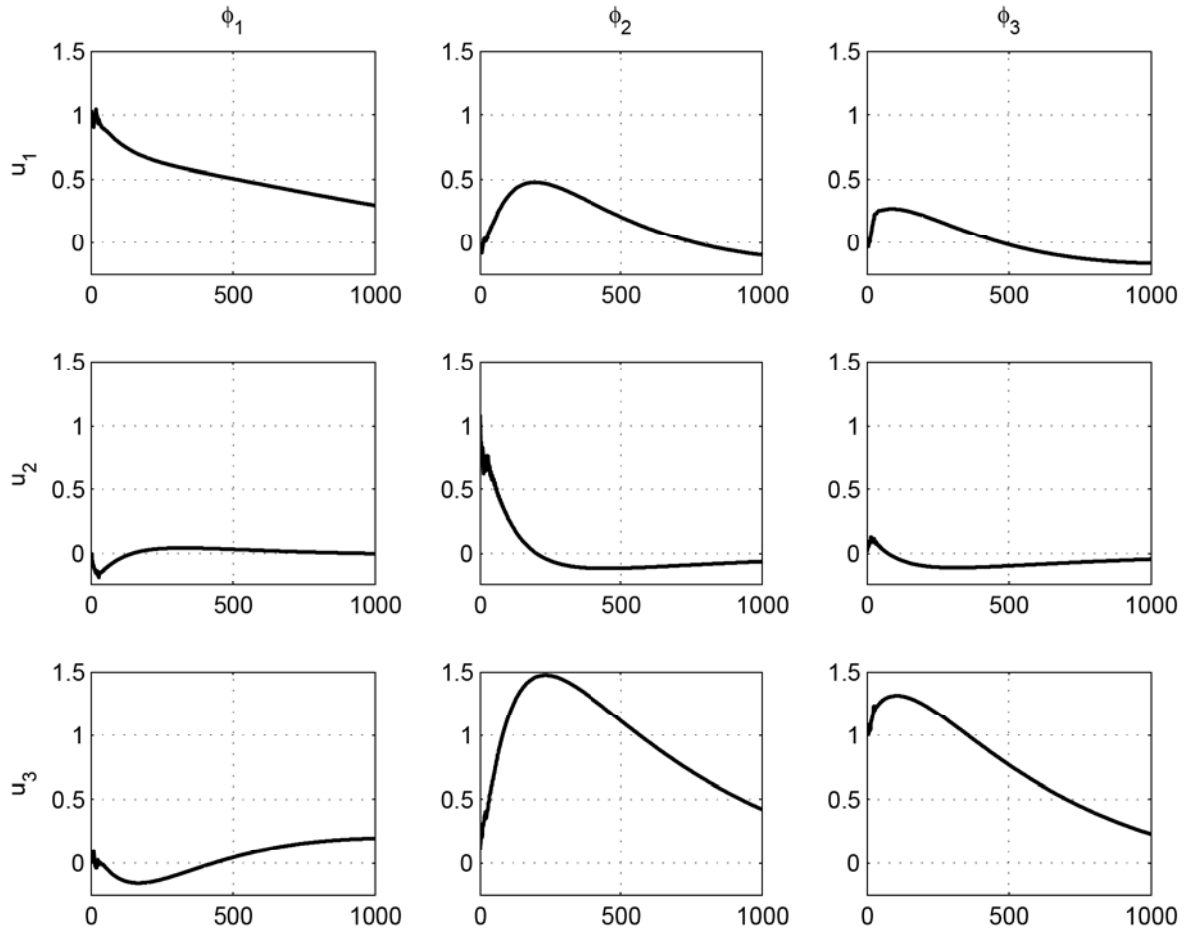


Figure 5: Impulse-response function of the factors (ϕ).

develops a 1.3 effect on the curvature factor, and dies out only slowly after that. It also increases the short factor dramatically. The maximum cumulative effect here is 1.5, indicating that a unit ϕ_3 -innovation develops a more than one unit effect on the short factor over time.

3.5. Impulse-response analysis in yield space

The innovations to these factors are of course uncorrelated by construction, but the VAR structure implies that the curvature factor is leading the short factor (because curvature innovations develop such a strong effect on the short factor over time). To get a better feel what these factor dynamics mean in terms of yield curve dynamics, we graph the impulse-response functions in yield space, see Figure 6. The top-left chart shows that an initial innovation to the long factor ϕ_1 . This essentially affects the long end, to a

lesser extent, the medium term yields, but leaves the short rate largely unaffected.

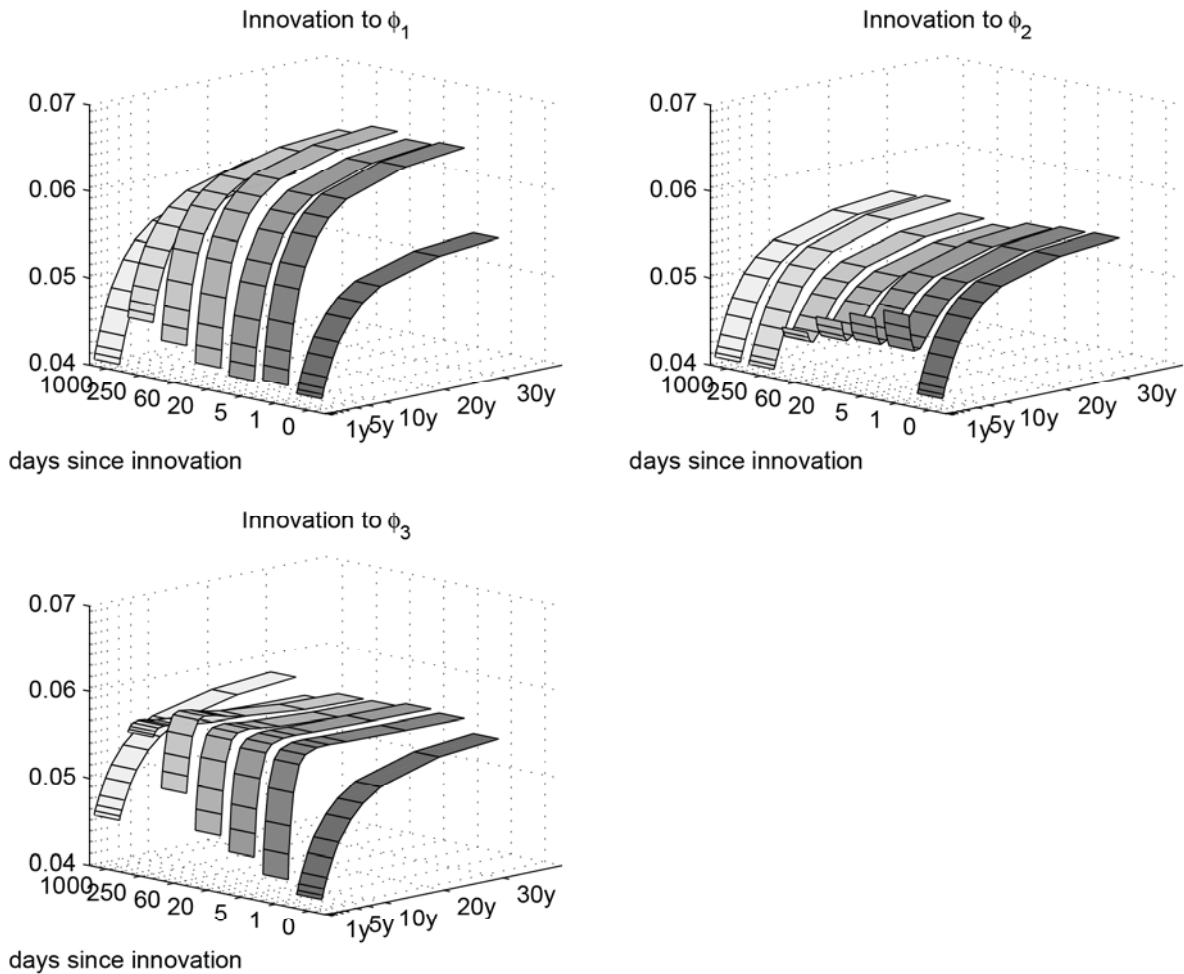


Figure 6: Impulse-response function in yield space.

An innovation to the curvature factor ϕ_3 (bottom-left chart) produces the most elaborate dynamics. This innovation of course increases curvature of the yield curve, and it normalizes only quite slowly. In the process of the returning to the steady state, the short rate also increases substantially, before it also returns to normal. During this process, the long rate first increases, then decreases, and then increases again.

An innovation to the short factor ϕ_2 (top-right chart) largely affects the short rate only. A strong enough positive short-factor innovation is able to produce a hump-shaped yield curve for a while. Also, through the VAR dynamics, such an innovation temporarily reduces the long end of the yield curve about one month after the innovation. Interestingly, this is quite compatible with Alan Greenspan's "conundrum." Between mid 2004 and end of 2005, the FOMC continuously raised the Federal Funds Rate, yet the

long rate steadily declined. This behavior seems to have puzzled the Fed, as becomes evident from chairman Greenspan’s testimony of February 2005,⁷ where he called this behavior of the term structure a “conundrum.” The impulse-response function in yield space of an innovation to the short factor suggests that the “conundrum” is really more the rule than the exception. It is exactly these dynamics — raising the short rate reduces the long rate — that the data impose on the factor model.

3.6. Variance decomposition

In order to assess the contribution of the factor innovations to the variance of yields at different maturities, we extend the usual approach from VAR analysis. Inverting the structural VAR model (14) yields structural the vector moving-average (VMA) representation

$$\phi_t = \bar{C} + C_0 u_t + C_1 u_{t-1} + C_2 u_{t-2} + \dots \quad (26)$$

Truncating this infinite VMA at a finite (but long) lag length, allows us to approximate the factors ϕ_t in terms of current and lagged structural innovations. In particular, we can decompose each factor into the contributions from the individual innovations. Define

$$\phi_t^i = C_0^i u_t^i + C_1^i u_{t-1}^i + C_2^i u_{t-2}^i + \dots + C_s^i u_{t-s}^i, \quad (27)$$

where C_j^i denotes the i^{th} column of C_j and u_t^i the innovation to factor i . The decomposition of the factors into the contributions from each innovation

$$\phi_t = \bar{C} + \phi_t^1 + \phi_t^2 + \phi_t^3 \quad (28)$$

carries over to the structural factor model of the yield curve,

$$r_t = k\bar{C} + k\phi_t^1 + k\phi_t^2 + k\phi_t^3 + \varepsilon_t. \quad (29)$$

Notice that the first four components on the right-hand side are approximately uncorrelated, because the orthogonality of the factor innovations u carries over to ϕ^i (27).⁸ This is not true for the idiosyncratic component ε . However, ε is small (see Figure 3), and therefore contributes only little to the variance of r . Defining $r_t^i = k\phi_t^i$ leads to the variance decomposition of the yield curve,

$$\text{var}(r_t) = \text{var}(r_t^1) + \text{var}(r_t^2) + \text{var}(r_t^3) + \text{error}, \quad (30)$$

⁷The Federal Reserve Board’s Semiannual Monetary Policy Report to the Congress before the Committee on Banking, Housing, and Urban Affairs, U.S. Senate, February 16, 2005.

⁸Of course, the uncorrelatedness of ϕ^i is not perfect because it depends on the truncation lag of the VMA representation of the factors. We find that setting $s = 2500$ yields acceptable results.

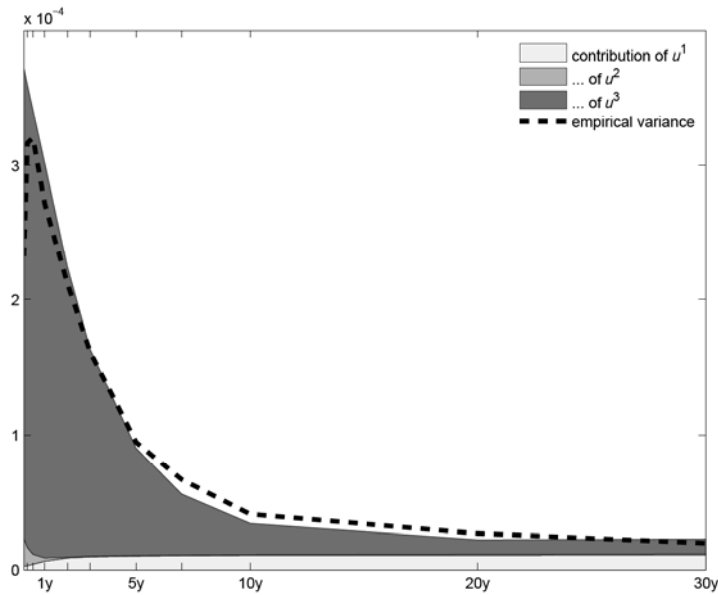


Figure 7: Relative importance of factor innovations for the variance of interest rates.

where the error in the above equation is small.

Figure 7 depicts the variance decomposition of the yield curve. The total height of the shaded area represents the contribution of the structural innovations; the relative contribution of each innovation is distinguished by different grayscales. The comparison with the empirical variance of the yield curve (dashed line) shows that the approximation error is small except for the 1- and 3-month maturities. Our sample contains no observations at the 1-month maturity before August 2001, when interest rates were exceptionally high. Therefore, the estimate of the empirical volatility at the 1-month maturity is biased downwards.

Over all, the innovations of the curvature factor are the main drivers of the volatilities of interest rates of all maturities. They account for more than 90 percent of the variance up to the 3-year yield, at longer maturities the importance of the curvature innovations declines but remains above 50 percent. The innovations to the long factor are negligible for the variance of short maturity yields. Their relative importance increases towards the longer end: long innovations explain about 45 percent of the variance of the 30-year yield. Finally, innovations to the short factor have almost no influence on the volatilities of the interest rates. They contribute about 5 percent to the variance of the shortest and longest yields. For intermediate maturities, they contribute almost nothing.

3.7. Results using the Diebold-Li data

As a robustness test, we repeat the estimation with the data that was used by Diebold and Li (2006).⁹ They use monthly data from January 1985 to December 2000 with constant maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months. The yields are extracted from market prices using the Fama-Bliss method (“unsmoothed Fama-Bliss yields”).

Repeating the computations with this alternative data set reveals strikingly little differences. We find that a VAR(1) is sufficient to remove serial correlation of the residuals. This is similar to the the 30 lags that are necessary with daily observations. The model yields an average RMSE of 6.6 bp, and a maximum RMSE of 18 bp (compared to 10 bp and 35 bp, respectively, using daily Treasury data). The impulse-response functions in yield space look very similar as well. The only significant difference is in the values of α . Using the monthly Diebold-Li data, (α_2, α_3) is estimated to be (0.0127, 0.6171), as opposed to (0.1133, 0.6798) with our main data set. Accordingly, the loadings are slightly different. The short factor diminishes somewhat faster to zero as maturity increases, and the loading of the curvature factor peaks around 1.25 years, as opposed to around 2.0 years with the main data set.

4. Where is the central bank?

The impulse-response analysis (particularly the one in yield space) invites us to think about the influence of monetary policy on the term structure of interest rates. It is true that the short rate is by and large controlled by the Federal Reserve. Indeed, our short factor is very close (but not exactly identical) to the Federal Funds Rate. As a first impulse, we might thus be tempted to identify monetary policy with the short factor. This interpretation would entail that the Fed is really quite powerless in affecting longer term yields.

However, the equalization of monetary policy with the short rate does not provide a complete picture of what central banks can do. The short rate is just the trading department of the central bank, so to speak. Yet, central banks influence the term structure also through other channels. For instance, a statement of a central bank governor might convey information that changes of the short rate are to be expected over the next few months. Such a statement — if it indeed contains information which was not already known by the market — would most probably affect yields with medium terms to maturity. In the parlance of our model, this would constitute a curvature innovation. As we have seen, innovations to the curvature translate into later movements of the short factor through the dynamics of the VAR. They also have a much stronger and longer lasting effect on the yield curve than pure short factor innovations have. Talking might be more important in monetary policy than trading.

⁹These data are available from Professor Diebold’s website, <http://www.ssc.upenn.edu/~fdiebold/papers/paper49/FBFITTED.txt>.

The central bank might even occasionally generate innovations into the long factor. For instance, a central bank may be able to change the perception of the public about how hawkishly it intends to fight inflation, or about its long term inflation target. Such a change of perception would certainly constitute an innovation to the long factor. Of course, the long factor should in theory be influenced by many other things as well, such as the real trend growth rate of the economy or the extent of long term uncertainty, but central bank strategy, its communicated targets, its perceived competence and credibility are also part of this factor.

We are thus led to the conclusion that the central bank is really present in all three factors. Its trading desk is well captured by the short factor, its communication about the intended course of action over the medium term is part of the curvature factor, and finally its target and credibility have a bearing on the long factor.

An illustration of this conclusion is given in Figure 8 which decomposes the 3-month, the 2- and the 10-year yields into the components driven by long, short, and curvature innovations. The first two columns show that the u-shaped development at the short end of the yield curve from 2001 to 2007 was essentially driven by curvature innovations. In addition, we observe that the component driven by curvature innovations leads the 3-month yield by a few months. This highlights the importance of the Fed's communication for movements of the short rate. Their decrease from the beginning of 2001 to mid 2004 and subsequent increase until mid 2006 was apparently well communicated in advance by the Fed and did not constitute unexpected movements (shocks) of the short rate at the time they occurred. In fact, innovations to the short factor played virtually no role after 2002, until the very recent events in the sub-prime crisis.

There are two noticeable episodes when the short factor innovations had a prominent effect: the first event is the attacks of September 11, 2001. The contribution of the short factor innovation to the short end of the term structure is clearly visible (see the top panel of Figure 9, which is an enlarged display of Figure 8). In the days following the attacks, the short factor was subject to innovations that lead to a sudden decrease of the 3-months interest rate by 50 basis points. The Fed had immediately started to offer liquidity on a large scale, and the Federal Funds Rate began to decline on September 13 (the first day of trading after the attacks). The Fed reduced the target rate of the Federal Funds Rate by 50 basis points just a few days later (on September 17, a Monday). The adjustment of the target rate was just a confirmation of the reduction of the Federal Funds Rate that had happened a few days before. Yet, it still appears that the reduction of the target was a surprise. This led to a further reduction of the short factor and of the short-end of the yield curve, but this second move was short-lived. The whole event was obviously not expected, and therefore not communicated beforehand by the Fed. Consequently, this move is not visible with a lead in the curvature factor, but is instead picked up in our model as an innovation to the short factor.

The second event is the recent 50-basis point discount rate cut that the Fed announced on August 17, 2007 "to promote the restoration of orderly conditions in financial markets" (from the Federal Reserve Press Release of that day), in the midst of the

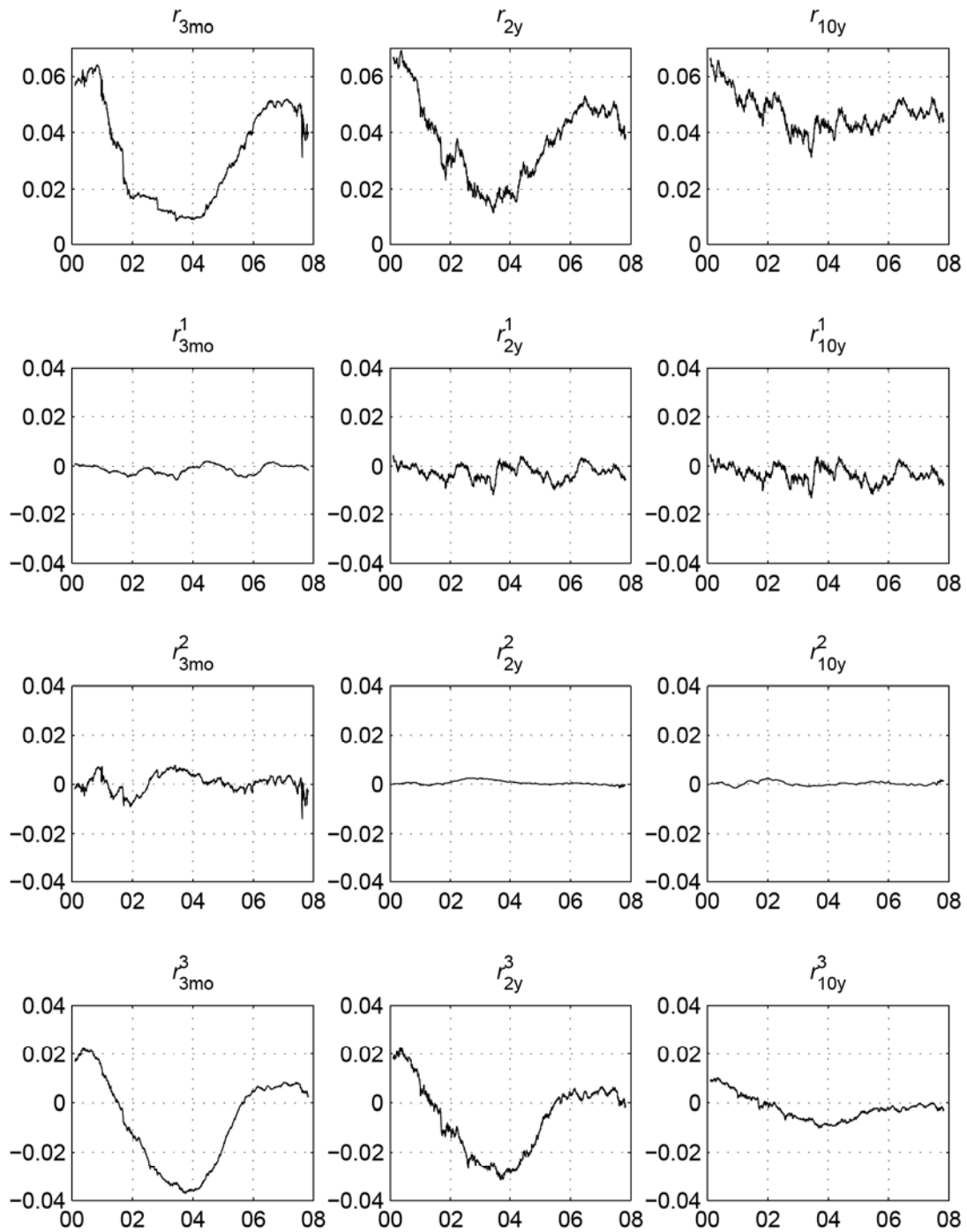


Figure 8: Historical decomposition of select yields into components driven by long, short, and curvature innovations.

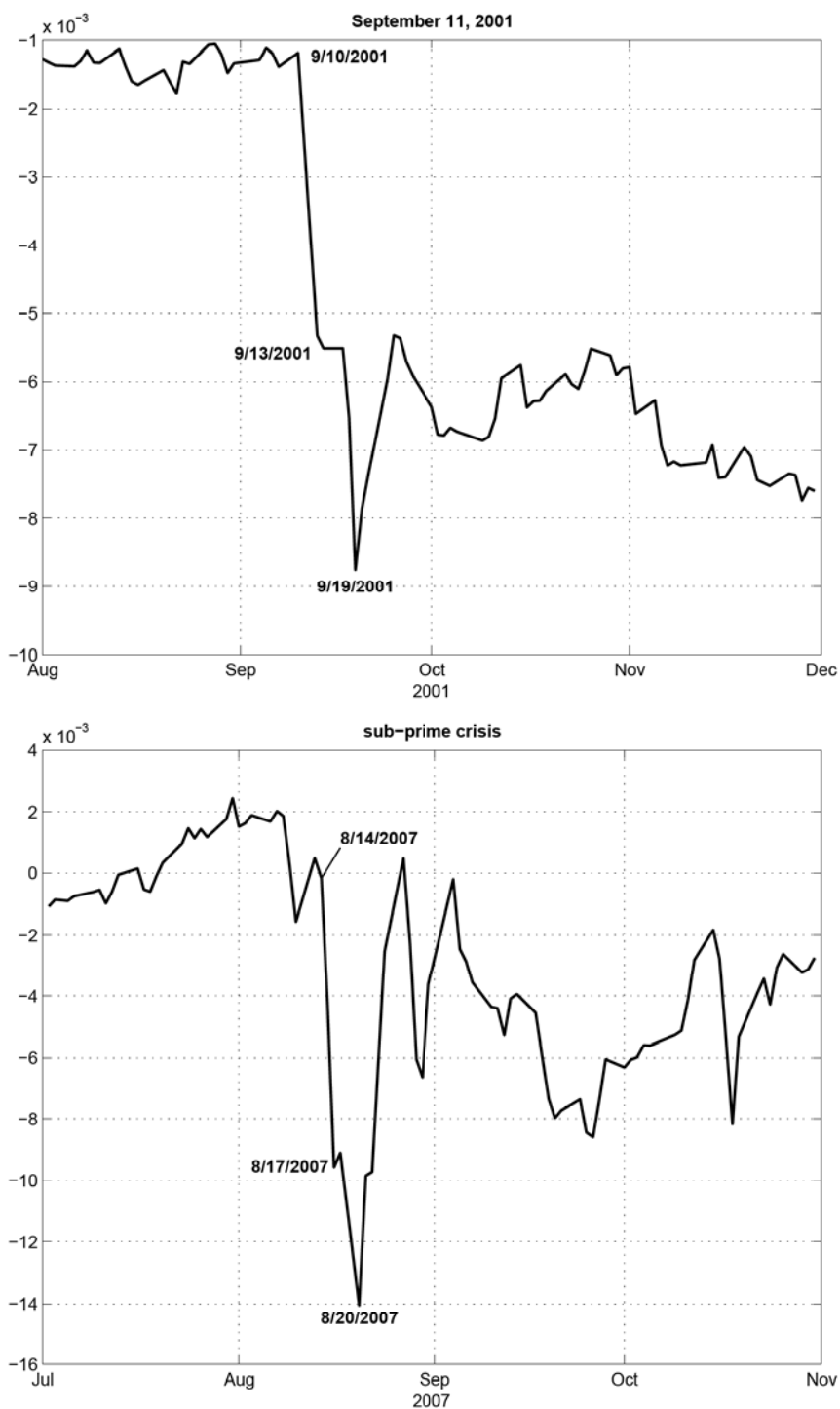


Figure 9: Contribution of short factor innovations to three-months yields. *Top*: Sudden negative impact after the attacks of 9/11. *Bottom*: Surprise action of the Fed (reduction of discount rate on August 17, 2007) in the course of the sub-prime crisis.

sub-prime crisis. This move seems to have been a surprise for most market participants because it is not detectable in the leading curvature factor. Instead, it is picked up in our model as an innovation to the short factor (the bottom panel of Figure 9 again depicts the effect of short rate innovations on the three-month interest rate). This innovation was very strongly negative, and led to an almost immediate drop of the three-month rate by 140 basis points — an impact much stronger than the one our model indicates after the September 11 catastrophe, and also much greater than the 50 basis point cut of the discount rate itself. However, unlike in the September 11 event, the effect was short-lived. In the days subsequent to the initial rate cut, the short factor experienced strong positive innovations which almost completely offset the initial effect on the short end of the yield curve within just ten days. In the following, and until the end of our sample on October 30, 2007, the short factor has experienced more volatility than what was usual before, and has therefore had a greater impact on the volatility of the short end of the yield curve. We interpret this as evidence that there is more volatility in Fed actions right now than there used to be.

5. Discussion

In this paper, we show how to construct intelligible factors that describe the term structure of interest rates and its dynamics. We believe that such factors are useful for practitioners because they provide a parsimonious description of the yield curve (the interpretability aspect), and also provide a language to discuss innovations to the yield curve, because innovations to one factor are independent of innovations to other factors (the orthogonality aspect). The orthogonality property is important: what does it mean to discuss a possible level shock in the Nelson-Siegel model (which does not have orthogonal innovations), if we know that a level shock typically comes together with a slope shock?

Our three factors do not only assume the interpretation as ‘long,’ ‘short,’ and ‘curvature’ that is imposed upon them by construction. We argue that they lend themselves also to interpretations related to monetary policy actions, due to the VAR structure of the factors. In this VAR, the curvature factor turns out to be a leading indicator of the short factor. We interpret this as saying that the curvature factor captures the communication of the Fed with regard to its future monetary policy stance. If the Fed indicates that it is likely to tighten monetary conditions, this induces market participants to expect higher interest rates in the medium term. In our model, such a communication is picked up as a positive curvature innovation. The short factor, on the other hand, captures surprise actions of the Fed. The fact that the short factor has only a minor effect on the term structure suggests that most monetary policy action is carefully communicated beforehand. We identify two exceptions to this rule, namely the September 11 shock, and the recent surprise reduction of the discount rate in the midst of the sub-prime crisis.

Beyond helping us to interpret current events, the VAR that is contained in our

model is also able to trace the dynamic effects of innovations to the three factors. This fact can be useful for several market participants. For instance, the model allows the central bank to quantify the immediate effects of its actions on the term structure, as well as the expected effects further down the road. Similarly, the model is potentially of value to bond portfolio managers. The model can be used to simulate the dynamics of the return distribution of bond portfolios, conditional on priors about factor innovations. This fact should be helpful to manage risk or optimize returns of a bond portfolio.

References

- ANG, A., AND M. PIAZZESI (2003): "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 50(4), 745–787.
- BLISS, R. R. (1997): "Testing Term Structure Estimation Methods," *Advances in Futures and Options Research*, 9, 197–231.
- COX, J. C., J. E. INGERSOLL, JR., AND S. A. ROSS (1981): "A Re-examination of Traditional Hypotheses about the Term Structure of Interest Rates," *Journal of Finance*, 36(4), 769–799.
- (1985): "A Theory of the Term Structure of Interest Rates," 53, 385–407.
- DAI, Q., AND K. SINGLETON (2000): "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943–1978.
- DE POOTER, M. (2007): "Examining the Nelson-Siegel Class of Term Structure Models: In-sample fit versus out-of-sample forecasting performance," Discussion Paper TI 2007–043/4, Tinbergen Institute, <http://www.tinbergen.nl/discussionpapers/07043.pdf>.
- DIEBOLD, F. X., AND C. LI (2006): "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130(2), 337–364.
- DUFFIE, D., AND R. KAN (1996): "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 379–406.
- LITTERMAN, R., AND J. SCHEINKMAN (1991): "Common Factors Affecting Bond Returns," *The Journal of Fixed Income*, pp. 54–61.
- MÖNCH, E. (2005): "Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach," Working Paper ECB WP544, European Central Bank, <http://www.ecb.int/pub/pdf/scpwps/ecbwp544.pdf>.

- NELSON, C. R., AND A. F. SIEGEL (1987): "Parsimonious Modeling of Yield Curves," *Journal of Business*, 60(4), 473–489.
- PYE, G. (1966): "A Markov Model of the Term Structure," *Quarterly Journal of Economics*, 80(1), 61–72.
- RUDEBUSCH, G. D., AND T. WU (2004): "A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy," Working Paper 2003–17, Federal Reserve Bank of San Francisco, <http://www.frbsf.org/publications/economics/papers/2003/wp03-17bk.pdf>.
- SVENSSON, L. E. (1994): "Estimating and Interpreting Forward Interest Rates: Sweden 1992–1994," NBER Working Paper 4871.
- VASICEK, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.

Swiss National Bank Working Papers published since 2004:

- 2004-1 Samuel Reynard: Financial Market Participation and the Apparent Instability of Money Demand
- 2004-2 Urs W. Birchler and Diana Hancock: What Does the Yield on Subordinated Bank Debt Measure?
- 2005-1 Hasan Bakhshi, Hashmat Khan and Barbara Rudolf: The Phillips curve under state-dependent pricing
- 2005-2 Andreas M. Fischer: On the Inadequacy of Newswire Reports for Empirical Research on Foreign Exchange Interventions
- 2006-1 Andreas M. Fischer: Measuring Income Elasticity for Swiss Money Demand: What do the Cantons say about Financial Innovation?
- 2006-2 Charlotte Christiansen and Angelo Ranaldo: Realized Bond-Stock Correlation: Macroeconomic Announcement Effects
- 2006-3 Martin Brown and Christian Zehnder: Credit Reporting, Relationship Banking, and Loan Repayment
- 2006-4 Hansjörg Lehmann and Michael Manz: The Exposure of Swiss Banks to Macroeconomic Shocks – an Empirical Investigation
- 2006-5 Katrin Assenmacher-Wesche and Stefan Gerlach: Money Growth, Output Gaps and Inflation at Low and High Frequency: Spectral Estimates for Switzerland
- 2006-6 Marlene Amstad and Andreas M. Fischer: Time-Varying Pass-Through from Import Prices to Consumer Prices: Evidence from an Event Study with Real-Time Data
- 2006-7 Samuel Reynard: Money and the Great Disinflation
- 2006-8 Urs W. Birchler and Matteo Facchinetti: Can bank supervisors rely on market data? A critical assessment from a Swiss perspective
- 2006-9 Petra Gerlach-Kristen: A Two-Pillar Phillips Curve for Switzerland
- 2006-10 Kevin J. Fox and Mathias Zurlinden: On Understanding Sources of Growth and Output Gaps for Switzerland
- 2006-11 Angelo Ranaldo: Intraday Market Dynamics Around Public Information Arrivals
- 2007-1 Andreas M. Fischer, Gulzina Isakova and Ulan Termechikov: Do FX traders in Bishkek have similar perceptions to their London colleagues? Survey evidence of market practitioners' views

- 2007-2 Ibrahim Chowdhury and Andreas Schabert: Federal Reserve Policy viewed through a Money Supply Lens
- 2007-3 Angelo Ranaldo: Segmentation and Time-of-Day Patterns in Foreign Exchange Markets
- 2007-4 Jürg M. Blum: Why 'Basel II' May Need a Leverage Ratio Restriction
- 2007-5 Samuel Reynard: Maintaining Low Inflation: Money, Interest Rates, and Policy Stance
- 2007-6 Rina Rosenblatt-Wisch: Loss Aversion in Aggregate Macroeconomic Time Series
- 2007-7 Martin Brown, Maria Rueda Maurer, Tamara Pak and Nurlanbek Tynaev: Banking Sector Reform and Interest Rates in Transition Economies: Bank-Level Evidence from Kyrgyzstan
- 2007-8 Hans-Jürg Büttler: An Orthogonal Polynomial Approach to Estimate the Term Structure of Interest Rates
- 2007-9 Raphael Auer: The Colonial Origins Of Comparative Development: Comment. A Solution to the Settler Mortality Debate
- 2007-10 Franziska Bignasca and Enzo Rossi: Applying the Hirose-Kamada filter to Swiss data: Output gap and exchange rate pass-through estimates
- 2007-11 Angelo Ranaldo and Enzo Rossi: The reaction of asset markets to Swiss National Bank communication
- 2007-12 Lukas Burkhard and Andreas M. Fischer: Communicating Policy Options at the Zero Bound
- 2007-13 Katrin Assenmacher-Wesche, Stefan Gerlach, and Toshitaka Sekine: Monetary Factors and Inflation in Japan
- 2007-14 Jean-Marc Natal and Nicolas Stoffels: Globalization, markups and the natural rate of interest
- 2007-15 Martin Brown, Tullio Jappelli and Marco Pagano: Information Sharing and Credit: Firm-Level Evidence from Transition Countries
- 2007-16 Andreas M. Fischer, Matthias Lutz and Manuel Wälti: Who Prices Locally? Survey Evidence of Swiss Exporters
- 2007-17 Angelo Ranaldo and Paul Söderlind: Safe Haven Currencies

2008-1 Martin Brown and Christian Zehnder: The Emergence of Information Sharing in Credit Markets

2008-2 Yvan Lengwiler and Carlos Lenz: Intelligible Factors for the Yield Curve

