

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

The impact of news on monetary policy expectations

### Permalink

<https://escholarship.org/uc/item/5ps194p0>

### Author

Bauer, Michael Dominic

### Publication Date

2010

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**The Impact of News on Monetary Policy Expectations**

A dissertation submitted in partial satisfaction of the  
requirements for the degree  
Doctor of Philosophy

in

Economics

by

Michael Dominic Bauer

Committee in charge:

Professor James D. Hamilton, Chair  
Professor Bruce N. Lehmann  
Professor Dimitris N. Politis  
Professor Irina A. Telyukova  
Professor Allan Timmermann

2010

Copyright  
Michael Dominic Bauer, 2010  
All rights reserved.

The dissertation of Michael Dominic Bauer is approved,  
and it is acceptable in quality and form for publication  
on microfilm and electronically:

---

---

---

---

---

---

Chair

University of California, San Diego

2010

## DEDICATION

To my beloved family, which made this endeavor possible.

EPIGRAPH

*No news is good news.*

—James Howell

## TABLE OF CONTENTS

Signature Page . . . . .	iii
Dedication . . . . .	iv
Epigraph . . . . .	v
Table of Contents . . . . .	vi
List of Figures . . . . .	ix
List of Tables . . . . .	x
Acknowledgements . . . . .	xi
Vita and Publications . . . . .	xii
Abstract of the Dissertation . . . . .	xiii
Chapter 1	
Revisions to Short Rate Expectations: Policy Surprises and Macroeconomic News . . . . .	1
1.1 Introduction . . . . .	2
1.2 Term structure model and estimation . . . . .	6
1.2.1 Risk-neutral dynamics . . . . .	6
1.2.2 Revisions to short rate expectations . . . . .	8
1.2.3 Physical distribution of revisions . . . . .	9
1.2.4 Money market futures . . . . .	10
1.2.5 Data and Estimation Method . . . . .	12
1.2.6 Parameter Estimates . . . . .	14
1.2.7 Reality check: Persistence of the federal funds rate	17
1.3 Term structures of volatility . . . . .	17
1.3.1 The effects of monetary policy . . . . .	20
1.3.2 Policy actions vs. macro news . . . . .	21
1.4 Comovement of forward rates . . . . .	22
1.4.1 Principal component analysis . . . . .	23
1.4.2 The regression approach reconsidered . . . . .	24
1.5 Measuring monetary policy surprises . . . . .	26
1.5.1 Specific examples of monetary policy surprises . . . . .	26
1.5.2 A horizon-specific measure for policy surprises . . . . .	27
1.6 The impact of macroeconomic data surprises . . . . .	29
1.7 Conclusion . . . . .	33
1.8 Appendix . . . . .	34
1.8.1 Revision to short rate expectations . . . . .	34

	1.8.2	Loadings of futures rate changes on factor shocks	35
	1.8.3	Additional evidence for policy inertia . . . . .	36
	1.8.4	Regression coefficients and $R^2$ implied by the model	37
	1.8.5	Inference about latent shocks and the revisions . .	37
Chapter 2		Term Premia and the News . . . . .	45
	2.1	Introduction . . . . .	46
	2.2	Dynamic Term Structure Model . . . . .	51
	2.2.1	Affine Gaussian DTSMs . . . . .	52
	2.2.2	Forward rates . . . . .	53
	2.2.3	The arbitrage-free Dynamic Nelson-Siegel model .	54
	2.2.4	Risk-neutral rates and forward risk premia . . . .	56
	2.2.5	Decomposing rate changes . . . . .	57
	2.2.6	Term Structure of Volatility . . . . .	59
	2.2.7	Eurodollar futures . . . . .	59
	2.3	Estimation of the unrestricted model . . . . .	61
	2.3.1	Econometric methodology . . . . .	61
	2.3.2	Parameter estimates . . . . .	64
	2.3.3	Term premium estimates . . . . .	66
	2.3.4	The impact of macroeconomic news . . . . .	70
	2.4	Restrictions on prices of risk . . . . .	73
	2.4.1	Identifying candidate specifications . . . . .	75
	2.4.2	Within-model simulation and posterior model prob- abilities . . . . .	78
	2.4.3	Economic implications of alternative models . . .	80
	2.4.4	Accounting for model uncertainty using reversible- jump MCMC . . . . .	82
	2.4.5	Forecast accuracy . . . . .	85
	2.4.6	Return predictability . . . . .	87
	2.5	Conclusion . . . . .	89
	2.6	Appendix . . . . .	90
	2.6.1	Change of measure . . . . .	90
	2.6.2	Basic MCMC algorithm and convergence diagnos- tics . . . . .	91
	2.6.3	MCMC algorithm: Latent indicator variables . . .	97
	2.6.4	Long-run revisions to short rate expectations . . .	100
	2.6.5	Reversible-jump Markov chain Monte Carlo . . .	100
Chapter 3		Inflation Expectations and the News . . . . .	110
	3.1	Introduction . . . . .	111
	3.2	Taylor Rules, Inflation Expectations and the News . . . .	113
	3.2.1	Taylor Rule evidence . . . . .	113
	3.2.2	Projecting a Taylor rule on past information sets .	116



3.3	TIPS Yields and Inflation Compensation . . . . .	118
3.3.1	Correlation between nominal rates and inflation compensation . . . . .	118
3.3.2	Responses to payroll surprises . . . . .	121
3.4	Survey-Based Measures of Inflation Expectations . . . . .	129
3.5	Projecting Future Inflation on Macro News . . . . .	132
3.6	Conclusion . . . . .	136

## LIST OF FIGURES

Figure 1.1:	Estimated means of futures rate changes . . . . .	14
Figure 1.2:	Estimated volatilities of pricing errors . . . . .	15
Figure 1.3:	Empirical vs. model-implied autocovariances of the short rate .	18
Figure 1.4:	Vol curves of money market futures in the four different news regimes . . . . .	19
Figure 1.5:	Empirical and model-implied results for the traditional regression approach . . . . .	40
Figure 1.6:	Policy surprises: Examples of revisions resulting from policy actions . . . . .	41
Figure 1.7:	Term structure of announcement effects . . . . .	44
Figure 2.1:	Average levels of actual, fitted and risk-neutral rates for unrestricted model . . . . .	67
Figure 2.2:	Time series of fitted rates and risk-neutral rates for unrestricted model . . . . .	68
Figure 2.3:	Effects of specific news events (unrestricted model) . . . . .	103
Figure 2.4:	Term structure of volatility implied by unrestricted model . . .	104
Figure 2.5:	Responses to macro news implied by unrestricted model . . . .	105
Figure 2.6:	Implications of unrestricted specification ( $M_1$ ) . . . . .	106
Figure 2.7:	Implications of favored specification ( $M_2$ ) . . . . .	107
Figure 2.8:	Comparison of alternative model specifications . . . . .	108
Figure 2.9:	Implications of Bayesian model averaging ( $BMA$ ) . . . . .	109
Figure 3.1:	Intraday responses of nominal rates, real rates and IC to payroll surprises . . . . .	125
Figure 3.2:	Responses of nominal rates and inflation compensation to non-farm payroll surprise . . . . .	127
Figure 3.3:	Revision of optimal inflation forecast following a payroll surprise	134

## LIST OF TABLES

Table 1.1:	Estimation results for benchmark model . . . . .	39
Table 1.2:	Regressions of rate changes on different measures of monetary policy surprises. . . . .	42
Table 1.3:	Effects of macroeconomic announcements . . . . .	43
Table 1.4:	Summary statistics for ratio of near-term rate changes . . . . .	43
Table 2.1:	Parameter estimates for unrestricted model . . . . .	65
Table 2.2:	Model specifications . . . . .	77
Table 2.3:	Forecast accuracy . . . . .	86
Table 2.4:	Return predictability . . . . .	88
Table 3.1:	Estimated univariate Taylor rules . . . . .	115
Table 3.2:	Correlations across subsamples . . . . .	120
Table 3.3:	Correlations across frequencies . . . . .	121
Table 3.4:	Intra-daily and daily responses to non-farm payroll surprise . .	123
Table 3.5:	Responses to payroll surprises . . . . .	128
Table 3.6:	Responses of survey-based inflation expectations to macro news	131

## ACKNOWLEDGEMENTS

I would like to thank Professor James Hamilton for his constant support as the chair of my committee. Through multiple drafts and numerous meetings, his guidance and advice has proved to be invaluable.

I am very grateful for the help and advice from my other committee members, Professor Bruce Lehmann, Professor Dimitris Politis, Professor Irina Telyukova, and Professor Allan Timmermann.

I would also like to acknowledge Professor Jonathan Wright and Professor Ioannis Ntzoufras, who provided particularly valuable feedback on my work.

This work benefited from the comments and insights of participants at research seminars at the Federal Reserve Board of Governors, the University of California, San Diego, the University of South Carolina, the Missouri Economics Conference 2009, the Northern Finance Association 2009 Annual Conference, the Spanish Economic Association 2009 Annual Conference, and the Eastern Economic Association 2010 Annual Conference.

## VITA AND PUBLICATIONS

2003	B.A. in Economics, University of St. Gallen, Switzerland
2005	M.A. in Quantitative Economics and Finance, University of St. Gallen, Switzerland
2005-2010	Teaching Assistant, University of California, San Diego
2009-2010	Lecturer, University of California, San Diego
2010	Ph.D. in Economics, University of California, San Diego

Bauer, Michael D., “Testing for Endogenous Growth”, *Master’s Thesis, University of St. Gallen, Switzerland*, VDM Publishing, 2009.

ABSTRACT OF THE DISSERTATION

**The Impact of News on Monetary Policy Expectations**

by

Michael Dominic Bauer

Doctor of Philosophy in Economics

University of California, San Diego, 2010

Professor James D. Hamilton, Chair

Expectations of future monetary policy are a crucial determinant of asset prices. This dissertation provides answers to the question how newly available information, such as news about the future stance of monetary policy or about the macroeconomic situation, affects these expectations. The first chapter characterizes how interest rates react to monetary policy actions and macroeconomic news. The second chapter decomposes interest rate changes into revisions of monetary policy expectations and changes in term premia, with the result that most high-frequency variation in interest rates is accounted for by expectations of future monetary policy. The third chapter targets the question what drives the revisions of nominal short rate expectations and documents an important role for inflation expectations.

# Chapter 1

## Revisions to Short Rate

## Expectations: Policy Surprises and Macroeconomic News

### Abstract

How do interest rates react to monetary policy actions and macroeconomic news? The conventional event study approach has several shortcomings, and this paper presents an alternative framework to answer this question, based on a dynamic term structure model that recognizes the heterogeneity of news events. My approach imposes no-arbitrage, parsimoniously captures the revisions to the entire expected short rate path, and integrates the analysis of different types of news. Policy actions are found to affect the entire yield curve, and the impact does not decline with maturity as suggested by previous studies. The impact of macroeconomic announcements reflects the fact that policy inertia plays an important role in how markets form expectations. Policy news lead to more varied effects than macro news, indicating that markets are surprised along more than one dimension by actions of the Fed.

## 1.1 Introduction

How do monetary policy actions affect the term structure of interest rates? Because of its high relevance to both market participants and policymakers this question has commanded considerable interest among researchers. The Fed controls the overnight interest rate, but the monetary transmission mechanism works through changes in interest rates at all maturities. Hence the effectiveness of monetary policy crucially depends on whether and how the Fed can impact interest rates other than the short rate. In a seminal paper Kuttner (2001) employs federal funds futures to measure monetary policy surprises. Using event study regressions he finds a significant effect of the surprises on yields at short and medium maturities, which however declines quickly with maturity. Other studies employing similar approaches come to the same conclusions (Poole and Rasche, 2000; Rigobon and Sack, 2004; Gürkaynak et al., 2005a; Hamilton, 2008): Policy surprises do affect interest rates, but the impact seems to decline with maturity.

A related literature analyzes how macroeconomic news affect the term structure, prominent contributions being Fleming and Remolona (1997), Balduzzi et al. (2001a) and Faust et al. (2007). These studies assess which macro announcements affect interest rates and what the sign and size of the responses are. The empirical approach is similar: Yield changes are regressed on surprise measures, and the equation is re-estimated for several maturities. These studies find that some macro news have an important impact on yields whereas others do not, and that the effects vary according to the maturity of the yields considered.

The regression approach employed in both strands of literature has some important shortcomings. Most importantly it does not uncover the effects of news events on the entire term structure, but only the effects on individual securities. The cross-sectional restrictions required by no-arbitrage are ignored. Imposing no-arbitrage is more attractive theoretically, but also entails important practical advantages such as improvements in statistical precision and the ability to predict responses of additional securities.

For the case of monetary policy analysis, the regression approach is particularly problematic. The surprise measure, which is derived from changes in



money market futures rates, and the dependent variable, usually the change in a bond yield, are both determined by the average change in the forward rate curve over a particular horizon. These regressions thus simply estimate the comovement between changes in forward rates at different maturities. Since the short rate naturally has a transitory component, the finding of decreasing explanatory power in the yield regressions of Kuttner (2001) and others is not surprising. It does not tell us anything substantial about the effects of monetary policy on the term structure.

Another shortcoming of the existing literature is that so far there has been no integrated analysis of the effects of both policy and macro news on the term structure of interest rates. The regression approach cannot be used to systematically compare different types of news in a common framework. Relevant questions are: What are the most important sources of volatility? How are rates across maturities affected by different types of news? What are the differences, if any, in the effects of policy actions and macroeconomic news?

This paper proposes a new way to study how policy actions and macro news affect interest rates. The object of interest is the *revision to the expected future path of the short rate* under the risk-neutral measure, since it captures the effects of a news event on the entire term structure. In order to parsimoniously capture the “revision”, which is an infinite-dimensional object, I employ a three-factor affine dynamic term structure model (DTSM). In addition to the advantages of imposing no-arbitrage, in particular the *reduction in dimensionality* that the cross-sectional restrictions achieve, this allows me to *integrate in a common framework* the news about monetary policy actions and the various kinds of macroeconomic news.

The key to integrating the different types of news is to explicitly account for the heterogeneity of these different sources of interest rate volatility. I achieve this by allowing the second moments of the model to depend on the “news regime”, i.e. the type of the news event that occurs on a given day. This is a simple but effective way to assess and compare the differential impact of policy actions and different macro news on interest rates.

In this way the paper also makes a contribution to the term structure literature: The *conditional structure* of my DTSM allows to identify and describe the

different sources of interest rate volatility. News about monetary policy and about the economy are the main drivers of changes in interest rates, however existing DTSMs treat all trading days in the same way, for example when estimating the “vol curve”, the term structure of volatility.<sup>1</sup> My model provides separate estimates of the vol curve for different types of news events, which reveals interesting differences. The conditional character of the DTSM tells us what really moves the market, and in which way it moves the market.

The DTSM used in this paper differs from conventional models in a second way: Only the risk-neutral dynamics are made explicit, but no pricing Kernel is specified. The reason is that the model is used to capture the changes in interest rates, but is not required to decompose these changes into changes in risk premia and changes in physical expectations of future short rates. It is therefore unnecessary to make explicit the risk-adjustment. The risk-neutral dynamics are specified so that the factors can be identified as level, slope and curvature, in the spirit of Christensen et al. (2007), which is convenient for estimation and interpretation of the results.

The paper shows that monetary policy generally has strong effects on the entire term structure. The volatility caused by policy actions reveals that long rates move just as much as short rates. Importantly the revisions show various different shapes: Some actions of the FOMC only move the short end of the yield curve and barely have an impact on longer rates, some have a hump-shaped impact, yet others leave the short end unchanged and move only long rates. My findings show that the impact of monetary policy does not decline with maturity as suggested by previous studies (Kuttner, 2001; Gürkaynak et al., 2005b,a), but instead that this impact strongly depends on the individual policy event, on average is hump-shaped, and causes significant movements in long rates.

A key result is that there is significant heterogeneity between different sources of interest rate volatility. The hypothesis of equal second moments on days with policy actions and on days with different types of macro news is strongly

---

<sup>1</sup>This also holds for regime-switching models such as Bansal and Zhou (2002) and Monfort and Pegoraro (2007), since they do not condition on observable information, i.e. do not distinguish between trading days.

rejected. More specifically the differences are the following: First, on days with macro news releases the vol curve is steeper at the short end and more back-loaded than on policy days. This indicates that markets expect the Fed to only sluggishly adjust the short rate in response to new information and constitutes evidence of policy inertia. Second, among the different types of news I consider, new employment reports are by and far the most important source of interest rate volatility. Third, revisions show much stronger comovement across horizons on macro news days than on policy day. This makes intuitive sense because on days with macro releases there is only one piece of new information, the data surprise. On policy days, on the other hand, there are several pieces of news – the current target choice and the information in the FOMC statement – which independently affect the market’s short rate expectations.

The fact that revisions in response to policy events come in various shapes parallels the findings of Gürkaynak et al. (2005b) who use principal component analysis to show that more than one factor is needed to describe monetary policy actions. Based on my model we can parsimoniously capture what happens on a policy day to the entire term structure, namely by describing the revision to the entire expected short rate path caused by the policy action. This is an improvement upon the target and path factors that Gürkaynak et al. (2005b) use to describe policy events, since it incorporates no-arbitrage and thus enables us to predict changes in yields and forward rates at any maturity. Based on this insight I develop a horizon-specific policy surprise measure and show its empirical success in predicting changes in the yield curve for U.S. treasury securities from near-term money market futures.

Finally the model provides a convenient and theoretically appealing framework for estimating the impact of macroeconomic announcements on the terms structure. Importantly, my estimates of the “term structure of announcement effects” are consistent with no-arbitrage. The key empirical findings: First, the case for policy inertia is strong. And second, the hypothesis of no response of far-ahead forward rates, which the DTSM allows me to test, is rejected for most announcements, supporting the “excess-sensitivity” evidence documented in Gürkaynak et

al. (2005a).

The paper is structured as follows: The model is introduced and estimated in Section 1.2. Section 1.3 presents estimates of the term structure of volatility conditional on the type of news event. Section 1.4 assesses comovement of rate changes in response to news events. In Section 1.5, after illustrating some specific instances of policy actions, I develop a new measure for monetary policy surprises and document its empirical success. Section 1.6 estimates the effects of macroeconomic data surprises on the term structure. Section 1.7 concludes.

## 1.2 Term structure model and estimation

### 1.2.1 Risk-neutral dynamics

Denote the rate for an overnight default-free loan between days  $t$  and  $t + 1$ , the short rate, by  $r_t$ .<sup>2</sup> It is assumed to be determined by three latent factors:  $r_t = X_{1t} + X_{2t} + X_{3t}$ . Assuming absence of arbitrage implies that there exists a risk-neutral measure  $\mathbb{Q}$  that prices all assets. The factor dynamics under  $\mathbb{Q}$  are specified as follows:

$$\begin{aligned} X_{1t} &= X_{1,t-1} + \varepsilon_{1t}^{\mathbb{Q}} \\ X_{2t} &= \rho X_{2,t-1} + \varepsilon_{2t}^{\mathbb{Q}} \\ X_{3t} &= \theta_1 X_{3,t-1} + \theta_2 X_{3,t-2} + \varepsilon_{3t}^{\mathbb{Q}} \end{aligned}$$

$$\varepsilon_t^{\mathbb{Q}} \stackrel{\mathbb{Q}}{\sim} N(0, V_{r(t)}), \quad E^{\mathbb{Q}}(\varepsilon_r^{\mathbb{Q}} \varepsilon_s^{\mathbb{Q}'}) = 0, \quad r \neq s$$

where  $\varepsilon_t^{\mathbb{Q}} = (\varepsilon_{1t}^{\mathbb{Q}}, \varepsilon_{2t}^{\mathbb{Q}}, \varepsilon_{3t}^{\mathbb{Q}})'$  is a martingale difference sequence (m.d.s.) under  $\mathbb{Q}$ , and the parameters  $\rho$ ,  $\theta_1$  and  $\theta_2$  satisfy stationarity restrictions.<sup>3</sup> Conventional DTSMs usually include three factors, since these explain the vast majority of

---

<sup>2</sup>I abstract from the facts that the overnight rate in the U.S., the effective fed funds rate, deviates from the target set by the monetary authority, and that the target has a step-function character. Both simplifications are inconsequential since I do not include observations of the short rate – inference is based on observed futures rates, which corresponds to average forward rates over a month (fed funds futures) or a quarter (Eurodollar futures).

<sup>3</sup>For the AR(1) process the restriction is  $|\rho| < 1$ . For the AR(2) process stationarity requires  $|\theta_2| < 1$ ,  $\theta_2 + \theta_1 < 1$  and  $\theta_2 - \theta_1 < 1$ , see Marmol (1995). I also assume that the roots of the AR(2) process are real.

variation in bond yields (Litterman and Scheinkman, 1991; Balduzzi et al., 1996). The specification here implies that the three factors are a priori identified as level, slope and curvature, in the spirit of Christensen et al. (2007). The first factor, which follows a random walk, corresponds to a level factor since shocks change expected future short rates at all horizons by the same amount. Empirically, far-ahead forward rates show a lot of variability (Gürkaynak et al., 2005b), which suggests that the short rate should have a unit root under  $\mathbb{Q}$ , since otherwise the model would imply that these forward rates are close to constant. The second factor serves as a slope factor since the effect of a shock declines with the horizon. The third factor has a hump-shaped impulse-response function, provided that the roots are sufficiently close to one. As noted by Backus et al. (1999), yield dynamics are hump-shaped, which is also evident from the shape of the term structure of volatility (Piazzesi, 2005). Hump-shaped dynamics can be generated either by an AR(2) factor<sup>4</sup> or by having a “central tendency” structure, where one factor reverts to another (Balduzzi et al., 1998; Christensen et al., 2007). Compared to existing models, our factor dynamics are most similar to those of Christensen et al. (2007), with the differences that their model includes a central tendency and is in continuous time.

The key novelty in the above specification is the inclusion of observable variance regimes: The shocks  $\varepsilon_t^{\mathbb{Q}}$  have a time-varying variance-covariance matrix  $V_{r(t)}$ . The function  $r(t)$  maps the calendar day  $t$  into one of  $R$  different variance regimes, according to the type of news that take place on that day. Specifically, we will set  $R = 4$ , the four regimes being FOMC announcement days, BLS employment report days, CPI/PPI days, and days with new retail sales data.<sup>5</sup> This formalizes the idea that market participants know what type of news occur on each day, thus I speak of “observable variance regimes”. Note the difference between this approach and “regime-switching” term structure models such as the ones of

---

<sup>4</sup>Another example of a term structure model that includes an AR(2) factor is the one of Startz and Tsang (2007).

<sup>5</sup>This assumes that only one event takes place on a given day, whereas in reality some days have more than one major news events. However these days are few in number, hence this simplifying assumption is inconsequential. The use of intraday data is a way to improve the precision of the estimates, however I leave this to future work.

Bansal and Zhou (2002) and Monfort and Pegoraro (2007): Those models treat the state variable that determines the regime as unobservable, whereas in our context everybody can observe which type of news event takes place on a given day and thus knows the value of the state variable. This greatly simplifies modeling and estimation, and is an obvious modeling choice if the goal is to condition on the different sources of news.

## 1.2.2 Revisions to short rate expectations

The expected path of the short rate under  $\mathbb{Q}$  determines the entire term structure at a specific date. For the forward rate contracted at date  $t$  for a loan from  $t+n$  to  $t+n+1$  we have  $f_t^n = E_t^{\mathbb{Q}} r_{t+n}$ , up to Jensen inequality terms.<sup>6</sup> Yields, forward rates, and money market futures rates, which will be considered in more detail below, are simply averages of these one-day forward rates. New information that moves the term structure is captured by the revision to the expected short rate path under  $\mathbb{Q}$ , that is

$$\{(E_t^{\mathbb{Q}} - E_{t-1}^{\mathbb{Q}})r_{t+n}\}_{n=0}^{\infty},$$

which I simply call a “revision”. Intuitively, this corresponds to the change in the forward rate curve, since  $f_t^n - f_{t-1}^{n+1} = (E_t^{\mathbb{Q}} - E_{t-1}^{\mathbb{Q}})r_{t+n}$ . Changes in all interest rates are determined by the revision on that day, which incorporates both changes in short rate expectations under the physical measure and changes in risk premia.

The above specification of the  $\mathbb{Q}$ -dynamics leads to a simple closed-form solution for the revision:

$$(E_t^{\mathbb{Q}} - E_{t-1}^{\mathbb{Q}})r_{t+n} = \begin{cases} \varepsilon_{1t}^{\mathbb{Q}} + \rho^n \varepsilon_{2t}^{\mathbb{Q}} + \frac{\phi_1^{n+1} - \phi_2^{n+1}}{\phi_1 - \phi_2} \varepsilon_{3t}^{\mathbb{Q}} & \phi_1 \neq \phi_2 \\ \varepsilon_{1t}^{\mathbb{Q}} + \rho^n \varepsilon_{2t}^{\mathbb{Q}} + (1+n)\phi_1^n \varepsilon_{3t}^{\mathbb{Q}} & \phi_1 = \phi_2, \end{cases} \quad (1.1)$$

where  $\phi_1$  and  $\phi_2$  are the roots of the characteristic equation of the AR(2) process. The derivation of these expressions is given in Appendix 1.8.1. We see that a shock

---

<sup>6</sup>We can safely ignore these since we exclusively consider *daily changes* in interest rates, e.g.  $f_t^n - f_{t-1}^{n+1}$ . Thus we ignore only the difference in the Jensen inequality terms for maturities  $n$  and  $n+1$ , which is negligibly small. Furthermore for money market futures rates there are no Jensen inequality terms, since their payoffs are linear functions of future short rates.

to  $X_{1t}$  leads to a parallel shift in the term structure, that shocks to  $X_{2t}$  die out exponentially, and that shocks to  $X_{3t}$  lead to a hump-shaped revision. Thus the shocks are naturally labeled level, slope and curvature shock, respectively.

### 1.2.3 Physical distribution of revisions

The revision is a linear combination of the risk-neutral innovations:

$$(E_t^{\mathbb{Q}} - E_{t-1}^{\mathbb{Q}})r_{t+n} = b'_n \varepsilon_t^{\mathbb{Q}},$$

where  $b_n$  denotes the loadings on the factor shocks for the revision at horizon  $n$ , given in equation 1.1. Thus under  $\mathbb{Q}$ , revisions and thus rate changes are Gaussian m.d.s. with variance  $b'_n V_{r(t)} b_n$ . If we were to specify a pricing Kernel, this would pin down the change of measure and provide us with the physical distribution of rate changes. This paper is not concerned with identifying and estimating risk premia, thus we can abstain from choosing a pricing Kernel and from estimating the physical factor dynamics, which is statistically challenging (Kim and Orphanides, 2005). However in order to estimate the model, we still need to specify the distribution of revisions under the real-world, physical measure  $\mathbb{P}$ .

The presence of risk premia generally leads to predictability of rate changes. Defining the forward risk premium  $\Pi_t^n = (E_t^{\mathbb{Q}} - E_t)r_{t+n}$  we have

$$(E_t^{\mathbb{Q}} - E_{t-1}^{\mathbb{Q}})r_{t+n} = (E_t - E_{t-1})r_{t+n} + \Pi_t^n - \Pi_{t-1}^{n+1}.$$

Whereas the first component, the revision to  $\mathbb{P}$ -expectations, is a m.d.s. under  $\mathbb{P}$ , the change in forward risk premia,  $\Pi_t^n - \Pi_{t-1}^{n+1}$ , introduces drift and serial correlation.<sup>7</sup> In order to deal with the drift I allow the revisions to have non-zero mean, which can differ across maturities.

With regard to serial correlation, first note that the autocorrelation of daily changes in money market futures has been found to be very small and economically

---

<sup>7</sup>For the case of constant risk premia this term is a constant, possibly non-zero. If forward risk premia are on average increasing with maturity, which is intuitively and empirically plausible, we have  $E(\Pi_t^n - \Pi_{t-1}^{n+1}) < 0$ . In general we cannot assume that either mean or serial correlation are zero.

insignificant (Hamilton, 2007). More importantly though, the sample I use consists only of days with particular news events. Since between any two event days there are numerous days that are not included in the sample, rate changes in my sample do not exhibit any significant serial correlation (evidence not shown). Thus we can safely assume that rate changes are serially uncorrelated.<sup>8</sup>

Hence we obtain the following distributional properties of the revisions under  $\mathbb{P}$ :

$$(E_t^{\mathbb{Q}} - E_{t-1}^{\mathbb{Q}})r_{t+n} = a_n + b'_n \varepsilon_t$$

$$\varepsilon_t \sim N(0, V_{r(t)}), \quad E(\varepsilon_r \varepsilon_s') = 0, \quad r \neq s$$

Note that the physical innovations  $\varepsilon_t$  have the same variance-covariance matrix as the risk-neutral innovations, a consequence of the diffusion-invariance principle (Piazzesi, 2009). In sum, under the physical measure, revisions are Gaussian with mean  $a_n$ , no serial correlation and variance  $b'_n V_{r(t)} b_n$ .

## 1.2.4 Money market futures

This paper uses money market futures, specifically federal funds futures and Eurodollar futures, for parameter estimation and all subsequent empirical analysis. These instruments are very liquid and provide a detailed picture of the forward rate curve. News events are quickly reflected in futures rates, which is why these are often quoted in the financial press. Another advantage over using treasury yields is that we do not need to extract a zero curve and forward rates from observed bond prices, since rates for a fixed set of maturities are directly available.

Federal funds futures, which were introduced by the Chicago Board of Trade (CBOT) in October 1988, settle based on the average effective fed funds rate over the course of the contract month. Denote the futures rate at time  $t$  of the  $i$ -month-ahead contract by  $FF_t^{(i)}$ . Letting  $m(t)$  be the day of the month corresponding to calendar day  $t$ , and  $M$  the number of days in a month (for simplicity assumed to be 31), settlement is based on the average short rate from  $t + iM - m(t) + 1$  to  $t + (i + 1)M - m(t)$ , the settlement rate. The cost to enter the contract is zero

---

<sup>8</sup>Even if there was some serial correlation and our model was misspecified in this respect, we could still give a Quasi-Maximum-Likelihood interpretation to our estimates.



and the payoff is proportional to the difference between the futures rate and the settlement rate.<sup>9</sup> Hence the pricing equation is

$$0 = E_t^{\mathbb{Q}} \left( FF_t^{(i)} - \frac{1}{M} \sum_{n=iM-m(t)+1}^{(i+1)M-m(t)} r_{t+n} \right) \quad (1.2)$$

and the daily change in the futures rate is accordingly determined by the average revision over the relevant horizon:

$$\Delta FF_t^{(i)} = M^{-1} \sum_{n=iM-m(t)+1}^{(i+1)M-m(t)} (a_n + b'_n \varepsilon_t) = \mu_{FFi} + h_{FFi,t}' \varepsilon_t \quad (1.3)$$

where  $\mu_{FFi}$  and  $h_{FFi}$  are the averages of  $a_n$  and  $b_n$ , respectively. Note that  $\mu_{FFi}$  is a parameter to be estimated, whereas  $h_{FFi,t}$  is a vector of loadings determined by the parameters of the risk-neutral dynamics and depending on  $t$  through the day of the month.<sup>10</sup> An explicit expression for  $h_{FFi,t}$  is given in Appendix 1.8.2. I use the one- to six-month-ahead fed funds futures contracts, denoted by FF1 to FF6 – contracts of longer maturity are not sufficiently liquid.

Eurodollar futures show deep liquidity for contracts expiring several years in the future. These contracts settle based on the 3-month LIBOR rate on the settlement date, which is the last day of the relevant quarter.<sup>11</sup> Denote by  $ED_t^{(i)}$  the rate of the  $i$ -quarter-ahead Eurodollar futures contract.<sup>12</sup> Letting  $Q$  be the days in a quarter (taken to be equal to 91) and  $q(t)$  the day of the quarter for calendar day  $t$ , the pricing equation is  $0 = E_t^{\mathbb{Q}} \left( ED_t^{(i)} - L_{t+iQ-q(t)} \right)$ , where  $L_{t+iQ-q(t)}$  is the 3-month LIBOR rate at the end of the relevant quarter. If we abstract from the

<sup>9</sup>Note that we ignore the effect of marking-to-market, i.e. the fact that payments are made before settlement, however the evidence of Piazzesi and Swanson (2008) indicates that this effect is likely to be negligible in our context.

<sup>10</sup>Both means and loadings depend on the day of the month  $m(t)$ . For  $\mu_{FFi}$  it is safe to ignore this dependence since we simply want to capture the average rate change for each contract.

<sup>11</sup>For more details on Eurodollar futures please refer to [http://www.cmegroup.com/trading/interest-rates/stir/eurodollar\\_contract\\_specifications.html](http://www.cmegroup.com/trading/interest-rates/stir/eurodollar_contract_specifications.html) (accessed 09/15/09).

<sup>12</sup>To clarify my terminology: The “one-quarter-ahead” contract, ED1, is a bet on LIBOR at the end of the current quarter, which is determined by the average expected short rate over the next quarter. This next quarter should be understood as the settlement quarter, since it is the expected value of the short rate over the course of this quarter that matters for the payoff.

credit risk inherent in LIBOR loans<sup>13</sup>, we have  $L_t = \frac{1}{Q} \sum_{j=0}^{Q-1} E_t^Q r_{t+j}$ . Hence

$$0 = E_t^Q \left( ED_t^{(i)} - \frac{1}{Q} \sum_{n=iQ-q(t)}^{(i+1)Q-q(t)-1} E_t^Q r_{t+n} \right) \quad (1.4)$$

which closely parallels equation 1.2, thus the rate changes are given by essentially the same formula as for fed funds futures –  $M$  is replaced by  $Q$  and  $m(t)$  by  $q(t)$ .<sup>14</sup> I denote the mean rate change by  $\mu_{EDi}$  and the loading for this contract by  $h_{EDi,t}$ . The contracts I consider are the ones that settle on the last day of the current and the next 15 quarters, denoted by ED1 to ED16.

### 1.2.5 Data and Estimation Method

The sample contains days between October 1988 (when fed funds futures started trading) and June 2007 (just before the recent turmoil in financial markets began) that fall into one of four regimes. The first regime contains days with FOMC announcements<sup>15</sup>, and the other three regimes are BLS employment reports, CPI/PPI releases, and releases of retail sales numbers. Days that would fall into more than one category are excluded. This results in a sample with 799 days, with 148 days in the first, 215 days in the second, 316 days in the third, and 120 days in the fourth regime. I choose these specific four regimes in order to see how policy actions and news about the employment situation, about inflation, and about aggregate demand differ in their impact on the yield curve.

The model parameters can be estimated using Maximum Likelihood Estimation (MLE). The focus on rate changes and the assumed absence of serial correlation leads to a particularly simple estimation procedure. As is common in term structure model estimation I introduce idiosyncratic pricing errors, denoted

---

<sup>13</sup>The credit risk resulting from commitment to a specific counter-party for three months instead of rolling over daily loans at the fed funds rate is measured by the LIBOR-OIS spread. Before August 2007 this spread was small and little volatile, hence we can neglect credit risk in our analysis.

<sup>14</sup>There is the slight difference that the relevant short rate horizon for Eurodollar futures starts one day earlier, the reason being that this horizon starts at the end of the quarter preceding the settlement quarter, whereas for fed funds futures it starts at the first day of the settlement month.

<sup>15</sup>Until December 2004 these are identified by Gürkaynak et al. (2005a). For the remaining period we take the days of the FOMC press release.

by  $\eta_t$ , because otherwise a low-dimensional factor model predicts a singular covariance matrix for higher-dimensional data. Denote by  $Y_t$  the vector with daily changes in the futures rates, measured in basis points:  $Y_t = (\Delta FF_t^{(1)}, \dots, \Delta FF_t^{(6)}, \Delta ED_t^{(1)}, \dots, \Delta ED_t^{(16)})'$ . The number of measurements is thus  $m = 22$ . The empirical specification is

$$Y_t = \underbrace{\mu}_{(m \times 1)} + \underbrace{H_t'}_{(m \times 3)(3 \times 1)} \underbrace{\varepsilon_t}_{(3 \times 1)} + \underbrace{\eta_t}_{(m \times 1)}, \quad (1.5)$$

where  $\mu = (\mu_{FF1}, \dots, \mu_{FF6}, \mu_{ED1}, \dots, \mu_{ED16})'$ , and  $H_t = (h_{FF1,t}, \dots, h_{FF6,t}, h_{ED1,t}, \dots, h_{ED16,t})$ . The pricing errors are assumed to be a Gaussian vector m.d.s., independent of  $\varepsilon_t$ , with contemporaneous covariance matrix  $R$  which is diagonal. Under these assumptions  $Y_t$  is serially uncorrelated and multivariate normal with mean  $\mu$  and covariance matrix  $\Sigma_t = H_t' V_r(t) H_t + R$ . The log-likelihood function is

$$\mathcal{L} = \sum_{t=1}^T -\frac{1}{2} \{T \log(2\pi) + \log(|\Sigma_t|) + (Y_t - \mu)' \Sigma_t^{-1} (Y_t - \mu)\}.$$

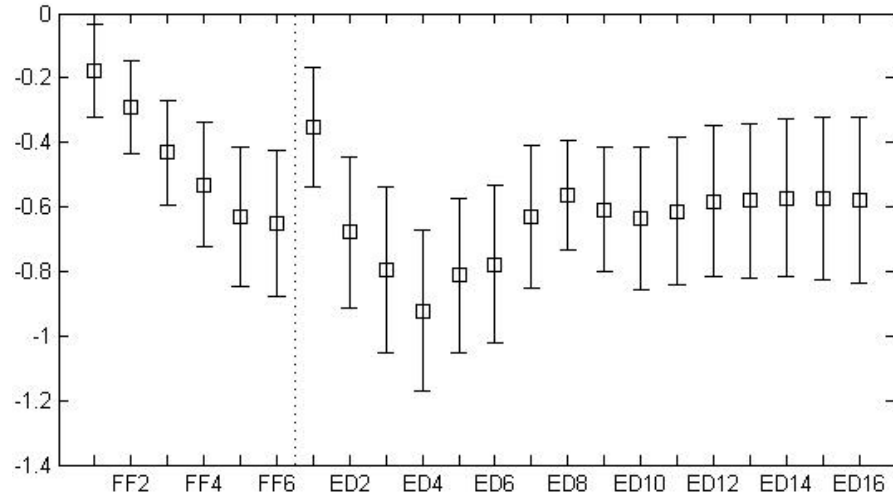
For the purpose of numerical optimization I reparameterize  $\mathcal{L}$  in order to ensure that the estimates are within the admissible parameter space: A Cholesky decomposition ensures positive definiteness of  $V_r$  in each regime. For the autoregressive root I take  $\rho = \lambda^2 / (1 + \lambda^2)$ , and similarly for  $\phi_1$  and  $\phi_2$ . For the volatilities of the pricing errors I let  $\sigma_{\eta,i} = e^{\zeta_i}$ .

The four shock-covariance matrices  $V_1$  to  $V_4$  each have six unique elements, which amounts to 24 parameters. The other parameters to estimate are  $\phi_1$ ,  $\phi_2$ , and  $\rho$ , as well as the 22 error variances and 22 means. The benchmark specification imposes  $\phi_1 = \phi_2 = \rho$  and thus has 69 free parameters. The large number of parameters leads to a significant but manageable computational burden when numerically maximizing  $\mathcal{L}$ . Optimization is performed using simplex and a gradient-based algorithms in turn. I try several different starting values and each time reach the same global optimum. Thus, despite the large number of parameters, MLE is easily feasible. This stands in contrast to the problems that come with estimating the physical dynamics of a DTSM, which make it difficult to perform MLE in that context, as reported for example by Kim and Orphanides (2005), Duffee and Stanton (2008) and Duffee (2009). The fact that we focus on cross-sectional dynamics

(Q) and do not try to estimate dynamic properties of the short rate (P) makes estimation a lot easier.

### 1.2.6 Parameter Estimates

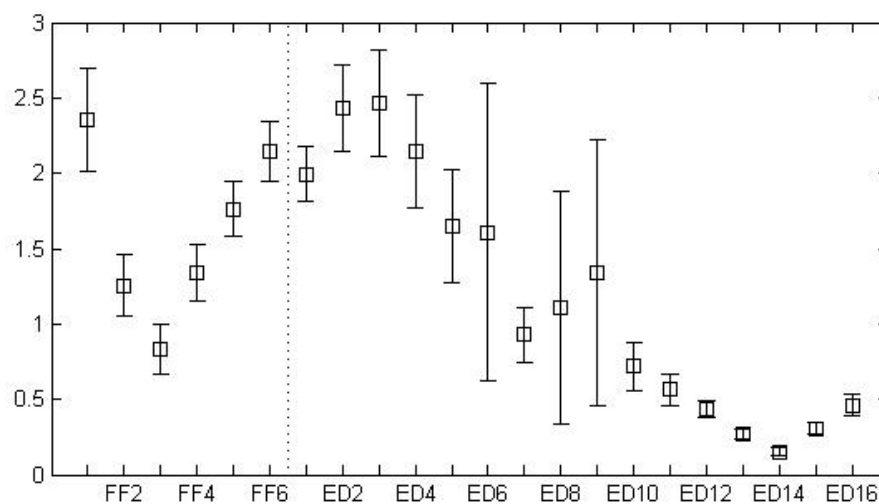
Table 1.1 and Figures 1.1 and 1.2 show the estimation results. The table reports the estimates for  $\rho$  and for the shock volatilities and correlations in each of the four regimes. It also reports the energy contents of the three principal components for each shock covariance matrix  $V_r$ , as well as the log-likelihood values for the benchmark version of the model and for more and less restricted versions. I report in parentheses robust Quasi-Maximum-Likelihood standard errors as suggested by White (1982), obtained using numerical approximations for gradient and Hessian.



**Figure 1.1:** Estimated means of futures rate changes

Estimated means of rate changes of money market futures, together with 95%-confidence intervals based on Quasi-Maximum-Likelihood standard errors, for benchmark specification of the model.

The shocks show important differences in variability and comovement across regimes. The shock variances are highest on days with a new employment report,



**Figure 1.2:** Estimated volatilities of pricing errors

Estimated volatilities of pricing errors together with 95%-confidence intervals based on Quasi-Maximum-Likelihood standard errors, for benchmark specification of the model.

and lowest on days with a new CPI or PPI report. Thus the release of a new employment report seems to have a bigger impact on short rate expectations than any other type of news event. Section 1.3 will consider vol curves and visualize the differences between regimes in terms of variability.

With regard to differences in comovement, the correlation between the shocks is generally higher on days with economic news than on policy days. This becomes particularly clear when we decompose the covariance matrix in each regime into its principal components (PCs) and show the energy content of each one: The first PC on policy days accounts for about 85% of the variance of the shocks, whereas for days with macro news the first PC accounts for 92-97% of the variability – factor shocks caused by macro news show stronger co-movement than shocks in response to policy news. Hence the revisions across maturities are more strongly correlated on news days, meaning that on days with macro news revisions tend to have more similar shapes than on days with policy actions. Section 1.4 will go into more detail about the differences in comovement that are implied by

these estimates.

Are the differences between the news regimes statistically significant? In order to test the hypothesis  $V_1 = V_2 = V_3 = V_4$  the model is re-estimated under this restriction (“equal  $V_r$ ’s”). Imposing this restriction leads to a significantly worse fit for the data. The log-likelihood is lower by about 187, which implies a likelihood-ratio statistic of 374. The number of restrictions and thus the degrees of freedom of the relevant Chi-square distribution is 18, leading to a minuscule p-value, hence we strongly reject the null of equal covariance matrices across regimes. The innovations to the term structure factors, i.e. the sources of interest rate volatility, have significantly different properties depending on the type of news causing them. In other words, the different sources of news show significant heterogeneity with regard to their impact on the term structure of interest rates.

For the AR(1) coefficient  $\rho$  we obtain an estimate of 0.9973. It is close to one because otherwise shocks to the transitory components would die out very quickly. The restriction that the roots of the AR(2) process,  $\phi_1$  and  $\phi_2$ , are equal to  $\rho$  is not rejected, as is evident from the log-likelihood value for the unrestricted version of the model (reported on the bottom as “different roots”).

Figure 1.1 shows estimates of the elements of  $\mu$ , the mean rate changes for each contract, together with 95%-confidence intervals. All are significantly negative, ranging from -0.2 to -1. This negative drift in forward risk premia makes sense based on the intuition that term premia are on average increasing with maturity, because rate changes imply a decrease in maturity by one day and thus a decrease in the average risk premium. The restriction that the mean rate change across futures contract is the same (“equal means”) leads to a likelihood-ratio statistic of 50 and hence is rejected. The restriction that  $\mu = 0$  (“zero means”) leads to a test statistic of 58 and is also rejected.<sup>16</sup>

Figure 1.2 shows point estimates and 95%-confidence intervals for the volatilities of the pricing errors. They generally decrease with maturity: The model fits the observed rate changes increasingly well for longer maturities.

---

<sup>16</sup>The 5%-critical values are 32.7 and 33.9, respectively, corresponding to degrees of freedom of 22 and 23.

### 1.2.7 Reality check: Persistence of the federal funds rate

Does the specification of the short rate as difference stationary make empirical sense? Do the parameter estimates imply dynamic properties for the short rate that accord with the evidence? As a reality check for the plausibility of model specification and estimation results I compare the model's implications with the empirical properties of the effective federal funds rate. The empirical autocovariance function of quarterly changes in the average fed funds rate, measured in basis points, is calculated over the sample horizon Oct-1988 to Jun-2007 – using quarterly averages avoids the problems that come with the discrete nature of changes in the target. The model-implied autocovariance function of quarterly changes in the average short rate is calculated by means of a simulation based on the model specification and estimates.<sup>17</sup>

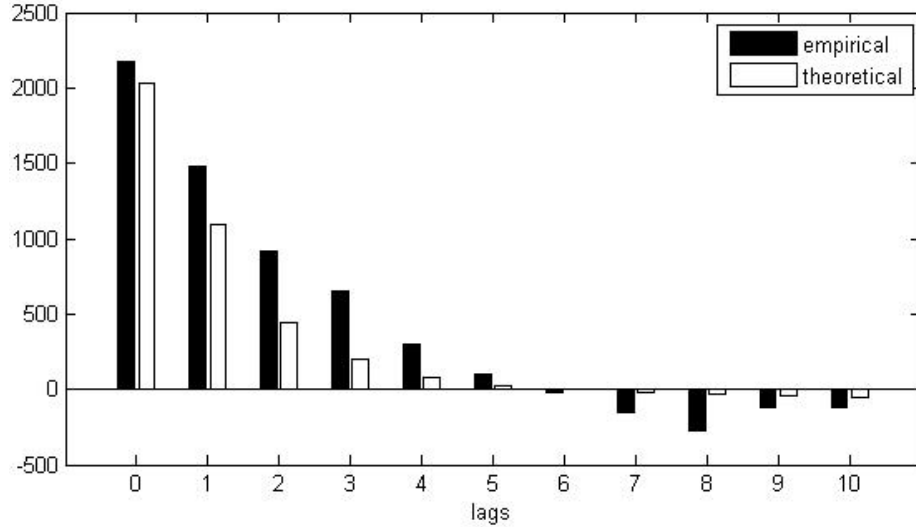
Figure 1.3 shows the result. It turns out that the model-implied variability and persistence properties correspond well to those of the actual short rate. The variance is roughly the same, around 2000. The first autocovariance is a bit higher in the data, but in the ball-park of the one implied by the model – both are between 1000 and 1500. Both autocovariance series remain positive until five lags and then turn negative. Since signs and even magnitudes of autocovariances are in accordance with data on the federal funds rate, my specification for the short rate and its model-implied properties based on estimates using money market futures appear to be plausible.

## 1.3 Term structures of volatility

The term structure of volatility, or “vol curve”, captures the volatility of rate changes across maturities. By allowing for heterogeneity in the sources of

---

<sup>17</sup>There are some caveats, neither of which undermine the usefulness of this exercise: First the model needs to be estimated on all trading days. The slight serial correlation of rate changes that was mentioned above is ignored in this estimation – its magnitude is very small and not economically significant anyways. Furthermore when re-estimating the model I only use one variance regime for simplicity, since here we only care about the average short rate dynamics across regimes. Finally the change of measure is ignored: The short rate process is simulated as if the risk-neutral and physical measure coincide.



**Figure 1.3:** Empirical vs. model-implied autocovariances of the short rate

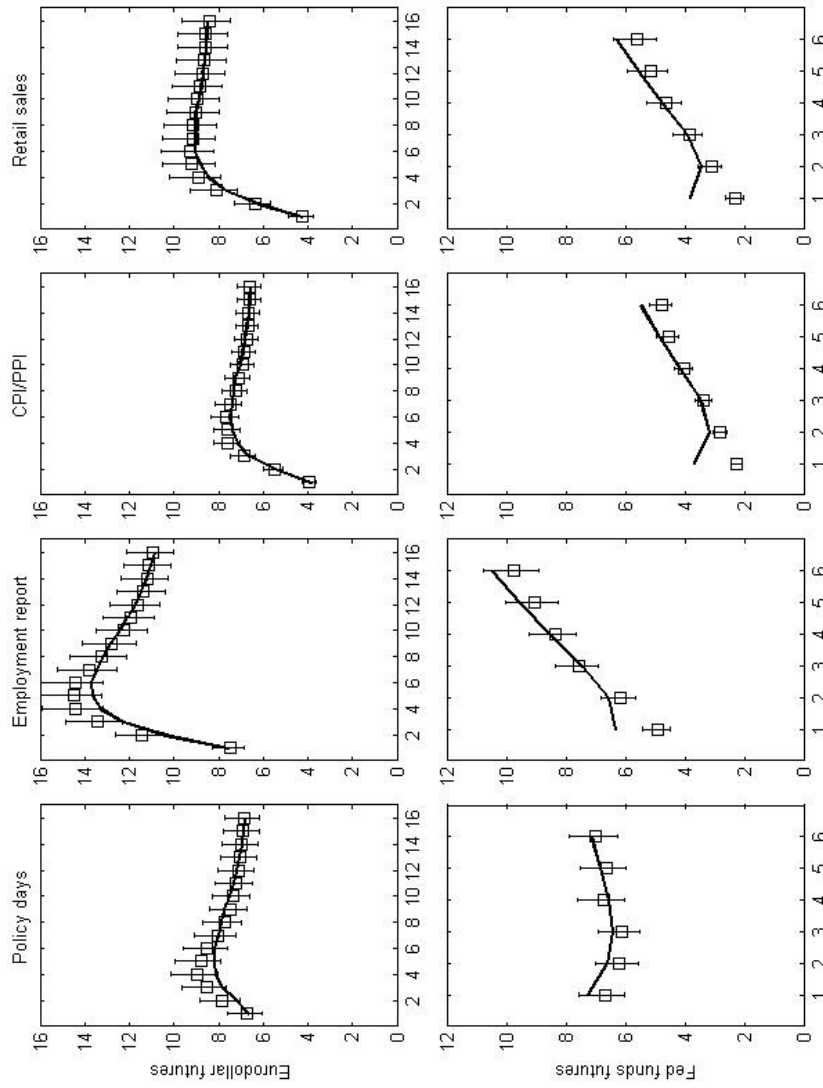
Autocovariances of changes in the average quarterly short rate. Empirical autocovariances are for the effective federal funds rate from Oct-1988 to Jun-2007. Theoretical autocovariances are based on the model-implied short rate process.

interest rate volatility, the model in this paper allows for a conditional assessment of the term structure of volatility: We can estimate different vol curves for each news regime, since the shock covariance matrix is allowed to vary depending on the news regime. The covariance matrix of futures rate changes is given by  $\Sigma_t = H_t V_{r(t)} H_t' + R$ . Denote the covariance matrix of futures rate changes in regime  $r$  by  $\Sigma_r$ .<sup>18</sup> The model-implied vol curve corresponds to the square root of  $\Sigma_r$ .

Figure 1.4 shows for each of the four different news regimes the empirical and model-implied volatilities of money market futures rates. The first row shows vol curves for Eurodollar futures and the second row for fed funds futures. Each of the four columns corresponds to a specific news event. The panels show sample standard deviations of daily rate changes in basis points for each contract, together with 95%-confidence intervals based on a Chi-square approximation. The thick black lines are the model-implied volatilities.

<sup>18</sup>Since the loadings  $H_t$  and thus  $\Sigma_t$  depend on the day of the month and on the day of the quarter, these need to be averaged out in order to obtain  $\Sigma_r$ .





**Figure 1.4:** Vol curves of money market futures in the four different news regimes

Empirical volatilities (sample standard deviations) with 95%-confidence intervals as well as model-implied volatilities for daily changes in Eurodollar futures rates and fed funds futures rates (in basis points) in each of the four news regimes.

The term structure model successfully captures the empirical volatilities: The shapes of the vol curves are closely replicated by the model, with model-implied vol curves generally within the confidence intervals and close to the point estimates of the empirical vol curves. The model specification allows enough flexibility to capture the hump shape of the vol curve (the back and the tail of the snake, see Piazzesi, 2001) and high volatilities of long forward rates (Gürkaynak et al., 2005b).

### 1.3.1 The effects of monetary policy

On the days in the first news regime policy actions by the Federal Reserve were the only major source of news. Interest rates moved when the FOMC's decision about the target or the content of the FOMC statement surprised market participants. The vol curves in the first column of Figure 1.4 visualize the variability of revisions caused by policy actions and thus inform us about the effects of monetary policy.

The Fed's actions clearly had a *significant impact on interest rates across the entire maturity spectrum*, which is evident from the large variability of revisions at all maturities on policy days. Particularly striking is that the Fed, by means of its actions and words, can affect the long end of the term structure as much as it can affect the short end: The longest Eurodollar futures contract (ED16) has a maturity of about four years and shows similar variability (about 7 basis points) as the fed funds futures contracts, which represent the short end of the vol curve. The impact of policy actions on short rate expectations is *strongest at a horizon of one to two years*, resulting in a distinct hump-shape.

These findings contrast with the conclusion suggested by Kuttner (2001) and Gürkaynak et al. (2005a) that the impact of monetary policy declines with maturity. These authors perform regressions of yield changes at different maturities on policy surprise measures, and find that regression coefficients and explanatory power ( $R^2$ ) decline with maturity. However, this regression approach does not describe the impact of policy actions on the term structure, but instead estimates the correlation of rate changes at different maturities. Section 1.4.2 will show that

my model reproduces these results, for which I will provide a new interpretation.

The bottom line is that Fed actions were highly effective in changing interest rates at all maturities. Since it is crucial for the transmission mechanism that the monetary authority can affect not only the overnight rate but also longer-term rates, the results indicate that monetary policy had potentially important effects on the economy.

### 1.3.2 Policy actions vs. macro news

Comparing interest rate volatility across news regimes, some important differences with regard to the level and the shape of the vol curves emerge. With regard to the level of volatility, futures rates are most volatile on days with a new employment report, evidenced by the high level of the vol curves in the second column. Evidently new information about the labor market is the biggest source of interest rate volatility, more important than news about monetary policy, the price level, or aggregate demand.

Considering the shape of the vol curves there is a striking difference between volatility caused by policy actions and volatility caused by macroeconomic news: On days with macroeconomic data releases the vol curves are steeply increasing at the short end and very back-loaded (the long end is at a higher level than the short end), resulting in a pronounced hump shape. On policy days however, futures rates at near-term and far-ahead horizons have similar volatility, and the hump shape is much less pronounced.

The shape of the vol curves corresponding to macro news constitutes evidence for policy inertia, the concept that changes in the stance of monetary policy are implemented by the Fed by slowly adjusting the target rate towards the new desired level. Intuitively the hump shape is related to policy inertia because evidently market participants revise their short rate expectations by much less over the coming months than over longer horizons, indicating that they expect the Fed to act slowly in response to the news. Furthermore Piazzesi (2001) showed how the back of the snake, i.e. the hump shape, can be attributed to policy inertia in the context of a term structure model that incorporates monetary policy. Thus

my evidence, the fact that macro news cause hump-shaped volatility, is distinctly in favor of policy inertia.

Notably my evidence stands in contrast to findings by Rudebusch (2006). We start from the same premise, namely that “changes in the path of expected future interest rates following the release of news about the state of the economy should reveal the degree of interest rate smoothing<sup>19</sup> because financial markets will expect an inertial central bank to distribute the policy rate changes over several periods” (p.26). The vol curves I estimate thus support the notion of policy inertia. Rudebusch however, based on a different empirical approach, concludes that the data speaks against policy inertia. Appendix 1.8.3 shows additional evidence in favor of policy inertia based on an approach more directly comparable to the one of Rudebusch and discusses possible reasons for the differences in our results.

The fact that vol curves on policy days are flatter and less hump-shaped than on days with macro news makes intuitive sense: On policy days, the Fed every so often surprises market participants with its choice of the target, as evidenced by Kuttner (2001), creating variability at the very short end of the term structure. This source of volatility pushes up the short end of the vol curve and accounts for the much flatter shape on policy days. My empirical analysis shows that the Fed not only creates volatility at the short end, but also has a significant impact on rates with longer maturities.

## 1.4 Comovement of forward rates

My framework allows us to compare the effects of news events also in terms of the comovement of rates at different maturities, since covariances are regime-dependent just like the vol curve. Stronger or weaker correlations across maturities tell us whether the revisions caused by a specific news event always look similar or whether they are rather varied in shape, and thus indicates whether there seem

---

<sup>19</sup>Rudebusch uses the terms “interest rate smoothing” and “policy inertia interchangeably”, whereas others, such as Piazzesi (2005) denote by “interest rate smoothing” the fact that the short rate is persistent, and use the term “policy inertia” to describe autocorrelated changes in the target federal funds rate, i.e. an inertial adjustment to the level desired by the Fed.

to be one or more independent sources of new information.

### 1.4.1 Principal component analysis

The comovement of the shocks to the term structure factors, measured by the off-diagonal elements of  $V_r$ , was shown in Section 1.2.6 to be stronger on days with macro news than on days with policy actions. The first PC accounts for the majority of variation in the shocks in the former case, whereas in the latter case the second PC contributes considerably to the variation.

Considering futures rates instead of shocks, note that the off-diagonal elements of the variance-covariance matrix of futures rate changes,  $\Sigma_r$ , measure the model-implied comovement of rates at different maturities, conditional on the news regime. A principal component analysis of this matrix<sup>20</sup> leads to the same conclusion as an analysis of the shock covariance matrix  $V_r$ . Stopping rules that help to determine the number of components describing common variance (Peres-Neto et al., 2005) imply that one component suffices on days with macro news, but that we need at least two components on policy days.<sup>21</sup>

The strong comovement of rate changes across the maturity spectrum implies that revisions caused by macro news usually come in similar shapes. Thus there seems to be one underlying source of volatility – a single factor is causing revisions. On the other hand the lower comovement on policy days indicates more varied shapes of revisions and thus independent sources of interest rate movements. This is consistent with the results of Gürkaynak et al. (2005a), who find that two factors are required to describe the variation in yields on days with policy actions.<sup>22</sup> However my analysis goes further in that it contrasts the effects of policy actions with the effects of macro news.

What causes the differences in comovement between regimes? If a particular news event usually has several pieces of new information which affect different parts

---

<sup>20</sup>I also performed a principal component analysis of the empirical covariance matrix of the futures rate changes, conditional on the same news regimes. The same conclusions applied.

<sup>21</sup>The results are omitted for sake of brevity but can be obtained from the author upon request.

<sup>22</sup>These authors use a PCA of the empirical covariance matrix of changes in money market futures.

of the term structure, then comovement in rate changes will be lower than for a news event with only one piece of information. On policy days markets learn the new target rate chosen by the Fed, as well as infer the Fed’s intentions about future policy from the FOMC statement. The target decision affects the short end of the term structure, whereas the information in the statement affects medium and long maturities. This can create a variety of possible revisions to the expected short rate path, as I will further exemplify in Section 1.5.1, leading to lower comovement than on days with macro news.

### 1.4.2 The regression approach reconsidered

The term structure model can be used to shed new light on the Kuttner-type regression approach common in the literature (Kuttner, 2001; Poole and Rasche, 2000; Gürkaynak et al., 2005b,a). I show that the model’s implications are consistent with the regression results. Furthermore I provide a re-interpretation of the regression approach and use it to get another perspective on the differences between news regimes.

Let’s consider regressions of rate changes in Eurodollar futures on a fed funds futures contract. I will use the FF3 contract, since it always has at least one FOMC meeting before delivery<sup>23</sup> and generally is a stronger measure of near-term policy surprises than the shortest contracts. This measure of the near-term policy surprise is closely related to the “Kuttner-shock”, which is a scaled change in the spot-month fed funds futures contract. As is common in the literature, I separately regress, for each Eurodollar futures contract, the futures rate change on the surprise measure.

Figure 1.5 shows, separately for each news regime and for each contract, the estimated response coefficients with 95%-confidence intervals based on White standard errors, together with model-implied response coefficients. Also shown are empirical and model-implied  $R^2$  for each regression. Appendix 1.8.4 shows the calculations underlying model-implied coefficients and  $R^2$ .

---

<sup>23</sup>Because the fed funds rate has a step-function character and only changes its level essentially every six weeks, the specification for the shortest fed funds futures contracts suffers from the problem that no rate change might occur until delivery.

The model-implied results closely correspond to the empirical results – model-implied regression coefficients are within the empirical confidence intervals, and model-implied and empirical  $R^2$  are very close to each other.<sup>24</sup> Thus, by capturing the second moments of revisions, the model in a sense encompasses the regression approach common in the literature.

The correlation between the near-term policy surprise and other rates always decreases with maturity. This makes intuitive sense: The surprise measure corresponds to the average revision of expected short rates over a specific horizon at the short end of the term structure. The dependent variables measure the average revision over longer horizons. Since the short rate has transitory components, revisions to expectations generally are less correlated the further apart the horizons are. Hence the finding of Kuttner-type regressions that explanatory power decreases with maturity are not surprising at all.

That  $R^2$  decreases more quickly for revisions caused by policy actions than for macro news reflects our previous finding of generally lower comovement on policy days.

With regard to the regression coefficients, an interesting difference between policy news and macro news emerges. In the case of policy surprises the regression coefficients are generally decreasing with maturity, whereas for macro news they show a distinct hump-shape. We saw above that vol curves to some degree are hump-shaped. For the policy news regime this hump shape is not reflected by the response coefficients. This indicates that the near-term policy surprise does not capture the effect of policy actions on the term structure, corresponding to our result that one factor is not enough to capture policy news. However, the near-term surprise measure apparently does a decent job in signaling the entire revision resulting from macro news.

---

<sup>24</sup>This holds for regressions not only for Eurodollar futures rates but also for yields (as in Kuttner, 2001) and forward rates (as in Gürkaynak et al., 2005b), as evidenced by additional calculations which I do not report here for sake of brevity.

## 1.5 Measuring monetary policy surprises

This section considers in more detail the revisions to short rate expectations that are caused by policy actions, i.e. monetary policy surprises.<sup>25</sup>

First the consideration of some specific days will show the variety of possible policy surprises. Then I undertake the task of predicting changes in long-term interest rates using short-term money market futures, which has been at the heart of numerous previous studies, using a new measure of policy surprises that naturally follows from the framework of this paper.

### 1.5.1 Specific examples of monetary policy surprises

To capture a monetary policy surprise we need to describe the revision to the expected short rate path. The model allows to parsimoniously describe the revision in terms of the values of the innovations to the term structure factors. Based on observed rate changes we can infer these factor shocks using a linear projection, as described in Appendix 1.8.5.

Figure 1.6 shows the impact of policy actions on four different dates. It shows actual changes in Eurodollar futures rates and fitted changes implied by the revision on that day. The dotted horizontal line indicates how the long-run expectation of the short rate has changed in response to the policy action. On all four days the target federal funds rate was increased by 25 basis points. On the dates shown in the top row, the FF1 rate increased by 10 and 8 bps respectively, whereas on the dates shown in the bottom row it did not change at all.

The revisions have very different shapes, indicating that the impact of monetary policy on the term structure differed significantly, which fits in with the evidence from above. Since Kuttner (2001) we know that the target rate change is not a useful measure for the policy surprise, and our graphs confirm this. Furthermore, the change in a near-term fed funds futures contract is not a good indication of what happens to the term structure either, evidenced by the differences between

---

<sup>25</sup>The difference between “policy surprise” and a “policy shock” is that while both are unanticipated, the latter is also exogenous. Clearly the changes in interest rates caused by policy actions are endogenous to the current economic situation.



the left and right column. This exemplifies the conclusion of Gürkaynak et al. (2005a) and our evidence from the previous section.

The bottom line is that *the entire revision* needs to be considered in order to understand the effect of the policy action on a particular day. The strength of the model is that it provides a detailed yet parsimonious description of the policy surprise, imposing the absence of arbitrage. This is achieved by having an underlying factor model for the short rate and imposing that rate changes must be due to revisions to the expected short rate path.

### 1.5.2 A horizon-specific measure for policy surprises

Although we can describe the revision resulting from policy actions graphically, many situations will require a numeric measure that summarizes the policy surprise. The state of the art is the approach of Gürkaynak et al. (2005a), henceforth GSS, namely providing two numbers, the “target factor” and the “path factor”, which are derived from near-term money market futures. These are simply the first two principal components of the futures rate changes, rotated such that the target factor has a unit impact on the nearest-term futures contract and that both factors have the same impact on the furthest-out futures contract. This approach has several shortcomings: First, by being simply a statistical summary of rate changes at the short end, it lacks the advantages of a term structure model. In particular, it does not imply changes of rates at arbitrary maturities that follow from no-arbitrage. Second, the measure is hard to interpret – it is not at all obvious what a specific number for the path factor means intuitively. Third, it leaves open the question of how we can construct a univariate summary of the policy surprise for a certain purpose.

The regression approach commonly employed in previous studies evaluates the predictive power of policy surprise measures, derived from near-term money market futures, for changes in yields and forward rates. There is a natural way to construct such a policy surprise measure based on the framework in this paper: First the entire revision to short rate expectations is inferred from changes in near-term money market futures. Then the predicted change in the relevant interest rate

can be calculated for any desired maturity. Importantly, this is possible because the no-arbitrage assumption allows us to predict changes in any security price that depends on the future path of the short rate. The predicted change can be interpreted as a horizon-specific policy surprise measure.

I construct this measure of monetary policy surprises for each day in the sample with a policy action. In order to back out the latent shocks and thus the revision I use the contracts FF1 to FF6 and ED1 to ED4. To compare the results with those of previous studies, I calculate the target and path factor of GSS, based on the same information set. The dependent variables are daily changes in the FF1 and ED4 futures contracts (to show the characteristics of target and path factor), changes in constant-maturity treasury bond yields at maturities two, five and ten years, as well as changes in six-month forward rates for loans maturing in two, five and ten years. Note that for each dependent variable a different horizon of the revision matters and thus a different surprise measure is constructed.

My sample contains the FOMC announcement dates from October 1988 to June 2007, excluding as before observations with employment reports, CPI/PPI news or new retail sales, and now also those that do not have yield data available. This leaves us with 148 observations. Data on yields and forward rates are those of Gürkaynak et al. (2006b).

Table 1.2 shows the results. Numbers in parentheses are White standard errors. The first three columns provide regression coefficients and  $R^2$  for regressions using only the target factor. The next four columns show the same for regressions using both target and path factors. These results are comparable to table 5 of GSS – differences result from the sample choice and the information set used to construct the factors. The first section of the table shows that in fact the target factor has a one-for-one impact on the one-month-ahead fed funds futures contract, that the factors are orthogonal, and that both factors have the same impact on the longest futures contract.

In the last three columns we see the predictive power of the horizon-specific policy surprise measure. First and foremost, the slope coefficients are all statistically significant at the 1% level, and larger than those on either target or path

factor. Treasury yields show a strong and significant response to policy surprises when the surprise is measured as I suggest. Furthermore the explanatory power of the univariate policy surprise measure that I constructed is about as large as that of target and path factors taken together: A univariate regression with my horizon-specific measure for policy surprises as explanatory variable explains the same amount of variation in yields and forward rates as a multivariate regression using the target and path factors of GSS. This results from the fact that my policy surprises are constructed specifically for the horizon under consideration.

The horizon-specific surprise measure proposed in this section can be used to predict changes in other securities, based on the fact that it captures the impact of the policy action on the expected short rate path. It does not only have intuitive appeal, but also is empirically successful when compared to target and path factors of GSS in terms of predictive power.

## 1.6 The impact of macroeconomic data surprises

Although it is well known that macroeconomic announcements have an important impact on interest rates, “few studies examine their impact on the yield curve as a whole,” as noted by Fleming and Remolona (1999b, p. 1). These authors fill this gap by estimating the impact of announcements on yields of different maturities, in order to estimate the “term structure of announcement effects.”<sup>26</sup> They find hump-shaped responses and attribute this to policy inertia. Gürkaynak et al. (2005b) perform a similar analysis, using forward rates instead of yields. The hump shape is clearly visible in their results as well, furthermore long forward rates react significantly to announcements. This is puzzling in the context of modern macro-models, and is thus termed “excess-sensitivity puzzle”.

Regressing rate changes for different maturities separately on the macro surprise measure is unsatisfactory because it does not impose absence of arbitrage.

---

<sup>26</sup>In the same paper the authors also develop a term structure model that incorporates the macro surprises and thus provides estimates of announcement effects that are consistent with no-arbitrage. My approach is fundamentally different in that no term structure factor is a priori identified with macro announcements.

This is not only a theoretical shortcoming, but also has two major practical disadvantages: First, without a term structure model that imposes no-arbitrage we cannot say anything about instruments other than the ones included in the regressions. Second, the separate regressions ignore cross-sectional restrictions which can help improve statistical precision.

My framework can provide estimates of the term structure of announcement effects based on a simple two-step procedure. The first step is to estimate the impact of a specific macro surprise, say a one standard deviation positive surprise in total payroll employment, on the latent factors. This is easily done by inferring the shocks for each day with macro news from futures rate changes (see Appendix 1.8.5) and regressing each of the three shocks on the macro surprises.<sup>27</sup> The coefficients on the surprise measure in each of the three equations tell us the values of the three shocks that are typically associated with the specific macro surprise. In the second step we use these values to calculate the revision that is implied by these shock values. This “typical” revision corresponds to the term structure of announcement effects. Importantly, it is consistent with the absence of arbitrage.

I use data on six different macroeconomic data releases: Non-farm payroll employment, the unemployment rate, hourly earnings, Core CPI and Core PPI (Bureau of Labor Statistics, BLS) as well as retail sales (Department of Commerce). The sample consists of all days with at least one data release between October 1988 and June 2007, which did not have a policy action, i.e. an FOMC announcement. The surprise component in the data release is calculated as the difference between the actually released number and the value expected by the market. To measure the market expectation I take the median market forecast, which is compiled by Money Market Services the Friday before the announcement.<sup>28</sup> The impact of the data surprises is estimated by including a constant and all six surprise measures on the right-hand-side of the regressions, thus if there are

---

<sup>27</sup>The regressions of all three shocks on the macro surprises constitute a system of Seemingly Unrelated Regressions (SUR). Since the explanatory variables are the same in each regression, equation-by-equation least squares is efficient in this case.

<sup>28</sup>Rigobon and Sack (2006) advance some points why this surprise measure might be contaminated with a considerable amount of noise. It might pay off to use intra-day data in this context, and possibly to address the measurement bias problem with new econometric tools. I relegate both issues to possible future work.

several news releases on one day, their impact is singled out by estimating partial effects.

Table 1.3 shows the numerical results, including the response coefficients for the futures contracts FF2 and ED4 as well as for the two-year yield. Numbers in parentheses are White standard errors. All announcements have a significant effect on the short end of the term structure, evidenced by the response of the FF2 contract. Also, all announcements with the exception of Core PPI lead to significant responses of ED4 and the two-year yield. Based on these securities alone, the evidence suggests that there is a significant impact of all of these macro news on the term structure.

A more detailed and accurate answer requires consideration of the term structure of announcement effects. This is captured numerically by the responses of the shocks, shown in the last three columns of Table 1.3. The response of  $\hat{\varepsilon}_{1t}$  measures the long-run response of short rate expectations, or “level impact” of the news, the response of  $\hat{\varepsilon}_{2t}$  measures the “slope impact”, and the response of  $\hat{\varepsilon}_{3t}$  measures the “curvature impact.” Figure 1.7 shows a graphical representation of these result: It plots the model-implied responses for all Eurodollar futures contracts, as well as the estimated long-run response. It also includes the estimated response coefficients and confidence intervals for separate regressions for each contract, corresponding to the conventional regression approach.

Comparing model-implied and unrestricted responses we see again that the model’s restrictions seem empirically plausible: The no-arbitrage-restricted responses closely correspond to the unrestricted estimates of the response coefficients.

One obvious pattern in the responses is a distinct hump-shape. Most announcements show this pattern, leading to increasing responses up to maturities of one to two years, with a decreasing response thereafter. A curvature shock  $\varepsilon_{3t}$  leads to a hump-shaped revision, hence we see a hump shape in the term structure of announcement effects in those cases where this shock responds significantly to the data surprise.

Increasing response coefficients at the short end are evidence for policy

inertia: The Fed is expected to act in response to the news, but not to immediately and fully adjust the target. Rather it will implement the new stance of monetary policy over a number of FOMC meetings, thus near-term responses are increasing in the horizon. The results here about the impact of macro news on short rate expectations add to the evidence on the term structures of volatility in Section 1.3 and make our case for policy inertia stronger.

Notably this analysis does not allow to distinguish between *intrinsic* and *extrinsic* policy inertia, as defined by Rudebusch (2006): The fact that markets expect a gradual response of the short rate to the macro news could be due either to an intentionally gradual adjustment on the part of the Fed (intrinsic inertia), or to persistence in the macroeconomic data (extrinsic inertia). If macroeconomic news is likely to be followed by news in the same direction, then even if the Fed immediately adjusts the fed funds rate to the new information, current macro news will predict subsequent changes in the short rate. In a recent study Hamilton et al. (2009) construct forecast of macro variables based on the news releases and in this way find some evidence for a deliberately measured pace of target rate changes on the part of the Fed, i.e. for intrinsic policy inertia.

The other important question is whether announcements move the long end of the term structure. My model allows us to assess this systematically by considering whether the release has a significant level impact. As we see in the right-most column of Table 1.3 this is the case for all announcements but the unemployment rate and hourly earnings. These two releases only move short rate expectations up to medium maturities, then their impact dies out. The unrestricted responses in Figure 1.7 also give some indication about the long-run impact of a specific announcement, however my approach allows us to actually *test* whether there are long-run effects.

That most releases have a significant level impact is consistent with the excess-sensitivity puzzle of (Gürkaynak et al., 2005b). Our finding corroborates the evidence of these authors that forward rates do not revert to a natural rate, or differently put that the short rate under  $\mathbb{Q}$  is not mean-reverting.

## 1.7 Conclusion

This paper introduces a coherent framework to describe and understand the impact of news on the term structure. The key question is: What are the effects of monetary policy surprises and macroeconomic announcements on interest rates? The framework integrates different types of news allowing for heterogenous sources of volatility. It is based on characterizing the revisions to the expected short rate path under the risk neutral measure that are caused by different news events.

The take-aways are: (1) The conventional regression approach of separately estimating the impact of some surprise measure on the rate of each instrument is not the right apparatus to assess the impact of news on the term structure. (2) Monetary policy actions affect the entire term structure, with the strongest impact at medium maturities. (3) Different policy actions vary greatly in their impact on interest rates, with these differences intuitively resulting from the independent pieces of information that markets receive. When measuring policy surprises, we thus need to take into account the relevant horizon. (4) Macroeconomic announcements differ in their impact on short rate expectations, but most lead to a strongly hump-shaped response and a significant long-run revision. (5) The evidence is clearly in favor of policy inertia: Market participants revise their expectations of the short rate in accordance with the Fed sluggishly adjusting its policy rate. (6) The hypothesis that far-ahead forward rates do not move in response to macro announcements is rejected for almost all data releases.

A valuable extension to this paper would be the use of intraday data in order to improve the precision of the estimates. In particular for monetary policy surprises we would like to make sure that the revision we estimate is the one caused by the policy action, and that there are no other confounding impacts. Considering tight windows around the policy announcements would corroborate our results about the effects of monetary policy on the term structure.

Another important task is to disentangle changes in short rate expectations from changes in term premia. This provides answers to other policy-relevant questions: How much of the volatility of observed rate changes is due to changes in short rate expectations? In response to specific news events, how did market

participants revise their expectations of monetary policy? Do risk premia systematically respond to macro news, and if yes in which way? In Chapter 2 I perform this exercise, and I find that at high frequencies changes in forward rates are mainly due to changing short rate expectations. This substantiates that the observed rate changes in response to news events in fact do predominantly signal changes in the expected path of the policy rate.

## 1.8 Appendix

### 1.8.1 Revision to short rate expectations

We need to find  $(E_t^Q - E_{t-1}^Q)r_{t+n} = (E_t^Q - E_{t-1}^Q)(X_{1,t+n} + X_{2,t+n} + X_{3,t+n})$ .

For the level factor, which follows a random walk without drift, we simply have

$$(E_t^Q - E_{t-1}^Q)X_{1,t+n} = \varepsilon_{1t}^Q.$$

For the slope factor, which follows an AR(1) process, the moving-average (MA) representation is  $X_{2t} = \sum_{j=0}^{\infty} \rho^j \varepsilon_{2,t-j}^Q$ , hence we have

$$(E_t^Q - E_{t-1}^Q)X_{2,t+n} = \rho^n \varepsilon_{2t}^Q.$$

The curvature factor follows an AR(2) process,  $X_{3t} = \theta_1 X_{3,t-1} + \theta_2 X_{3,t-2} + \varepsilon_{3t}^Q$ , which we rewrite as  $(1 - \phi_1 L)(1 - \phi_2 L)X_{3t} = \varepsilon_{3t}^Q$ , where  $L$  is the Lag-operator. Denote by  $\phi_1$  and  $\phi_2$  the roots of the characteristic equation  $y^2 - \theta_1 y - \theta_2 = 0$ , which are related to the parameters by  $\theta_1 = \phi_1 + \phi_2$  and  $\theta_2 = -\phi_1 \phi_2$ . We want to find the MA-representation of  $X_{3t}$ ,

$$X_{3t} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{3,t-j}^Q,$$

and for the Wold-coefficients  $\psi_j$  we have the difference equation  $\psi_j = \theta_1 \psi_{j-1} + \theta_2 \psi_{j-2}$  with initial conditions  $\psi_1 = \theta_1$  and  $\psi_2 = \theta_1^2 + \theta_2$  (see Brockwell and Davis, 2006, p. 92). For the case that we have distinct real roots  $\phi_1$  and  $\phi_2$  the solution to this initial value problem is given by  $\psi_j = (\phi_1^{n+1} - \phi_2^{n+1})/(\phi_1 - \phi_2)$  so that we obtain

$$(E_t^Q - E_{t-1}^Q)X_{3,t+n} = \frac{\phi_1^{n+1} - \phi_2^{n+1}}{\phi_1 - \phi_2} \varepsilon_{3t}^Q.$$



In the case that the roots are real and equal we get  $\psi_j = (1 + n)\phi_1^n$  and thus

$$(E_t^Q - E_{t-1}^Q)X_{3,t+n} = (1 + n)\phi_1^n \varepsilon_{3t}^Q.$$

### 1.8.2 Loadings of futures rate changes on factor shocks

Consider the loadings on the physical innovations  $\varepsilon_t$  of the rate change in the  $i$ -month-ahead fed funds futures contract, denoted by  $h_{FFi,t} = (h_{FFi,t}^1, h_{FFi,t}^2, h_{FFi,t}^3)'$ . The loading on the level shock is of course unity, i.e.  $h_{FFi,t}^1 = 1$ . The loading on the slope shock is

$$\begin{aligned} h_{FFi,t}^2 &= M^{-1} \sum_{n=iM-m(t)+1}^{(i+1)M-m(t)} \rho^n \\ &= \frac{\rho^{iM-m(t)+1}(1 - \rho^M)}{M(1 - \rho)}. \end{aligned}$$

For the loading on the curvature shock – we consider here only the case of equal roots – the average we need to calculate is

$$h_{FFi,t}^3 = M^{-1} \sum_{n=iM-m(t)+1}^{(i+1)M-m(t)} (1 + n)\rho^n$$

In order to find this average first consider the well-known summation formula for the geometric progression,

$$\sum_{k=m}^n r^k = \frac{r^m - r^{n+1}}{1 - r},$$

and take the first derivative with respect to  $r$  to obtain

$$\sum_{k=m}^n k r^{k-1} = \frac{m r^{m-1} - (n+1)r^n}{1 - r} + \frac{r^m - r^{n+1}}{(1 - r)^2}.$$

This can be applied to our summation to yield

$$\begin{aligned} h_{FFi,t}^3 &= \frac{(iM - m(t) + 2)\rho^{iM-m(t)+1} - [(i+1)M - m(t) + 2]\rho^{(i+1)M-m(t)+1}}{M(1 - \rho)} \\ &\quad + \frac{\rho^{iM-m(t)+2}(1 - \rho^M)}{M(1 - \rho)^2} \end{aligned}$$

The loadings for Eurodollar futures follow by analogy – simply replace  $M$  by  $Q$ , replace  $m(t)$  by  $q(t)$ , and take the first and last period of the relevant horizon for the summations to be one day earlier than in the case for fed funds futures (see footnote 14).

### 1.8.3 Additional evidence for policy inertia

One of the pieces of evidence presented in Rudebusch (2006) against the hypothesis of policy inertia is based on the following approach: Rudebusch calculates the ratio of 3-by-3-month forward rate to 3-month yield, based on intraday data on U.S. Treasury securities, for days with either a new employment report or new CPI data. In the case of policy inertia, the forward rate should move more strongly than the yield, and hence the ratio should usually be above one. Since the median and mean of this ratio are essentially equal to one in his sample, he concludes that “the case of little or no inertia is the relevant one” (p.29).

The above table presents comparable evidence, using the rate changes in the Eurodollar futures contracts ED2 to ED4 relative to the change in the contract ED1. The median and mean of this ratio is in all subsamples far above one. Hence for the futures rate changes in my sample, the evidence is strongly in favor of policy inertia also using Rudebusch’s approach, indicating that policy inertia is a fact rather than fiction.

How can the differences in our evidence be explained? Adjusting the sample window to the one used by Rudebusch does not change the qualitative results. The use of intraday data as opposed to my daily data is no likely candidate explanation, since the rate variation on the days under consideration is certainly caused to a large extent by the news event. The features of the yield data is the most likely explanation: Possibly 3-month and 6-month U.S. treasury securities move much more in lock-step than indicated by the expectations hypothesis, because of phenomena such as flight-to-security or hedging motives. The use of money market futures appears more likely to deliver reliable results.

### 1.8.4 Regression coefficients and $R^2$ implied by the model

The population parameters for regressions of rate changes in Eurodollar futures contracts on rate changes in the FF3 contract are equal to the covariance between the two variables, divided by the variance of the fed funds futures rate changes. The coefficient of determination is simply the squared correlation. The model parameters imply both variances and covariances, and hence predict the regression coefficients and  $R^2$ . Importantly, these depend on the news regime through the factor covariance matrix  $V_r$ . In regime  $r$ , the model-implied regression coefficients and coefficients of determination for a regression of Eurodollar futures contract  $i$  on the contract FF3 are

$$\begin{aligned}\beta_i &= \frac{Cov(\Delta ED_t^{(i)}, \Delta FF_t^{(3)})}{Var(\Delta FF_t^{(3)})} = \frac{Cov(h'_{EDi}\varepsilon_t, h'_{FF3}\varepsilon_t)}{Var(h'_{FF3}\varepsilon_t + \eta_t^{FF3})} \\ &= \frac{h'_{EDi}V_r h_{FF3}}{h'_{FF3}V_r h_{FF3} + \sigma_{\eta,FF3}^2}, \text{ and} \\ R_i^2 &= \frac{(Cov(\Delta ED_t^{(i)}, \Delta FF_t^{(3)}))^2}{Var(\Delta ED_t^{(i)})Var(\Delta FF_t^{(3)})} \\ &= \frac{(h_{EDi}'V_r h_{FF3})^2}{(h_{FF3}'V_r h_{FF3} + \sigma_{\eta,FF3}^2)(h_{EDi}'V_r h_{EDi} + \sigma_{\eta,EDi}^2)}.\end{aligned}$$

Here  $i = 1, \dots, 16$ , and  $\eta_t^{FF3}$  stands for the pricing error of FF3, which has a variance of  $\sigma_{\eta,FF3}^2$ , the square of the third diagonal element of  $R$ . The pricing error variances for the Eurodollar futures  $\sigma_{\eta,EDi}^2$  correspond to the squares of the 7th to 22nd diagonal elements of  $R$ .

The above notation ignores that the loadings of the futures depend on the day of the month and the day of the quarter, and hence the regression coefficient is a different one for each combination of these. To obtain an unconditional regression coefficient, I simply average out the day of the month/quarter.

### 1.8.5 Inference about latent shocks and the revisions

Given estimates of the model parameters the values of the latent shocks and thus the entire revision can be inferred from observed rate changes. Optimal inference about the latent shocks implies finding the conditional expectation of

the shock vector given the data on day  $t$ ,  $E(\varepsilon_t|Y_t)$ . Because of the normality assumption it can be calculated from a linear projection:

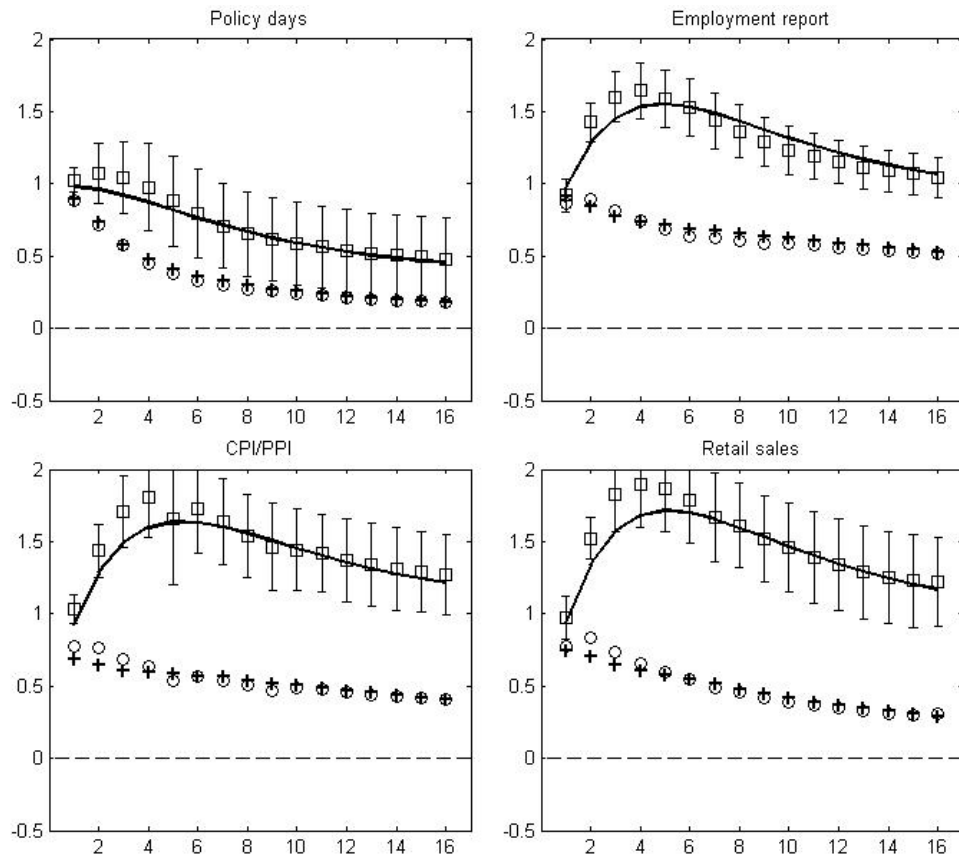
$$\begin{aligned}\hat{\varepsilon}_t = E(\varepsilon_t|Y_t) &= Cov(\varepsilon_t, Y_t)[Var(Y_t)]^{-1}(Y_t - \hat{\mu}) \\ &= V_{r(t)}H_t(H_t'V_{r(t)}H_t + R)^{-1}(Y_t - \hat{\mu})\end{aligned}\quad (1.6)$$

The fitted values for the futures rate changes are  $\hat{\mu} + H_t'\hat{\varepsilon}_t$ . The estimated revision to the  $\mathbb{Q}$ -expected short rate path is given by equation 1.1, where we substitute the estimated shocks for the unobserved errors. Note that we also have an estimate for the long-run revision  $\lim_{n \rightarrow \infty} (E_t^{\mathbb{Q}} - E_{t-1}^{\mathbb{Q}})r_{t+n} = \varepsilon_{1t}$ , which is simply equal to the level shock.

**Table 1.1:** Estimation results for benchmark model

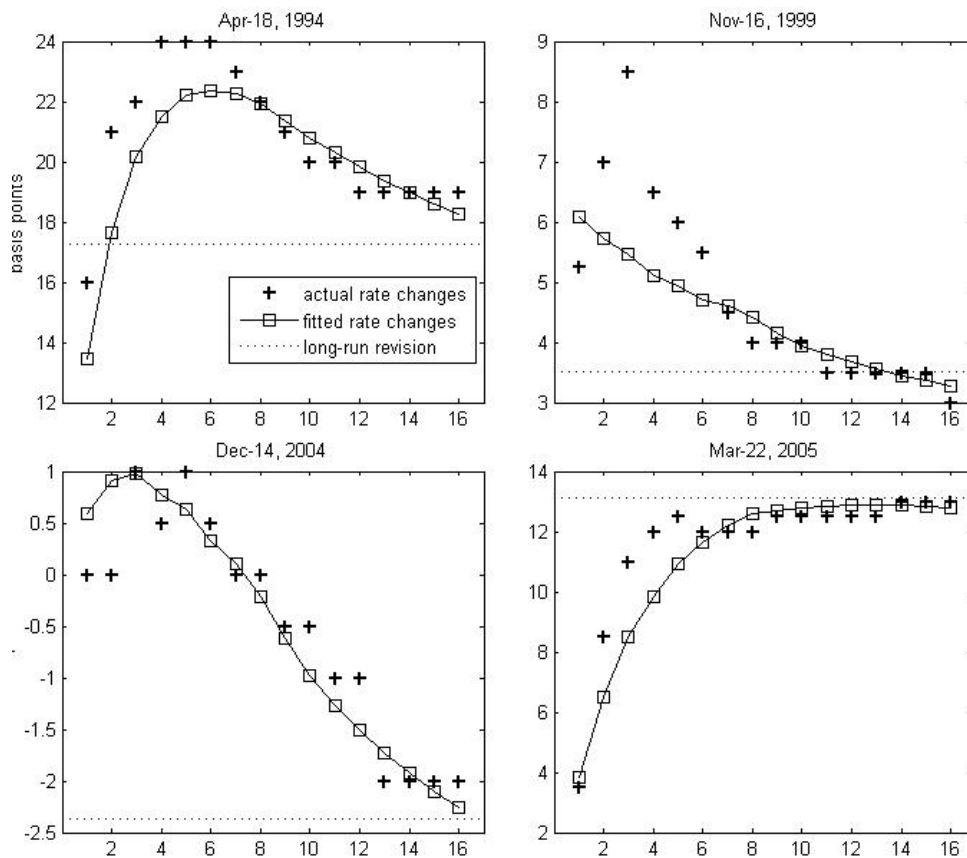
Parameter	policy days	empl. report	CPI/PPI	retail sales
$\sigma_{\varepsilon,1}$	6.5434 (0.4992)	10.0468 (0.6740)	6.3384 (0.3060)	8.2916 (0.9709)
$\sigma_{\varepsilon,2}$	9.9291 (1.3795)	11.3830 (0.9379)	7.2834 (0.3755)	9.9564 (1.3312)
$\sigma_{\varepsilon,3}$	0.0415 (0.0038)	0.0648 (0.0056)	0.0379 (0.0028)	0.0419 (0.0050)
$corr(\varepsilon_{1t}, \varepsilon_{2t})$	-0.6390 (0.0962)	-0.8437 (0.0301)	-0.8877 (0.0223)	-0.9394 (0.0201)
$corr(\varepsilon_{1t}, \varepsilon_{3t})$	+0.2676 (0.0883)	+0.5035 (0.0651)	+0.2358 (0.0565)	+0.2505 (0.0856)
$corr(\varepsilon_{2t}, \varepsilon_{3t})$	-0.3662 (0.0662)	-0.5426 (0.0752)	-0.4273 (0.0596)	-0.3312 (0.0949)
<i>Energy content</i>				
1st PC	85.37%	92.32%	94.50%	97.07%
2nd PC	14.63%	7.68%	5.50%	2.93%
3rd PC	< 0.01%	< 0.01%	< 0.01%	< 0.01%
$\rho$	0.9973 (0.0001)			
Log-likelihood	-31,422			
<i>Other models</i>				
different roots	-31,422			
equal $V_r$ 's	-31,609			
equal means	-31,447			
zero means	-31,451			

I use fed funds futures contracts FF1 to FF6 and Eurodollar futures contracts ED1 to ED16 to estimate the model. The sample consists of days with news events from Oct-1988 to Jun-2007. The number of days is 799, with 148, 215, 316, and 120 days in each of the four news regimes. Numbers in parentheses are robust standard errors. Also reported is the energy content of each principal components for each of the shock covariance matrices  $V^1$  to  $V^4$ , as well as the log-likelihood for the benchmark version of the model and for more and less restricted versions. For details please refer to text.



**Figure 1.5:** Empirical and model-implied results for the traditional regression approach

Regressions of changes in Eurodollar futures rates on changes in near-term federal funds futures rates (FF3) in different news regimes: Empirical response coefficients (squares) with 95% confidence intervals based on White standard errors, and model-implied response coefficients (thick line). Also shown are empirical coefficients of determination (circles) and the model-implied counterparts (crosses).



**Figure 1.6:** Policy surprises: Examples of revisions resulting from policy actions

Actual changes in Eurodollar futures rates (crosses) and fitted changes implied by revisions (squares) on four days with monetary policy actions. Also shown are the estimated long-run revisions (dotted lines). Note: The Fed increased the target by 25 bps on all of these days. On the dates shown in the top row, the FF1 rate increased by 10 and 8 bps respectively, whereas on the dates shown in the bottom row it did not change at all.

**Table 1.2:** Regressions of rate changes on different measures of monetary policy surprises.

	Target factor			Target and path factors			Horizon-specific surprise		
	const.	target	$R^2$	const.	target	path	const.	shock	$R^2$
<b>Futures</b>									
FF1	0.17 (0.15)	1.00* (0.03)	0.93	0.17 (0.14)	1.00* (0.03)	0.00 (0.01)	0.93		
ED4	-0.64 (0.60)	0.59* (0.16)	0.18	-0.09 (0.09)	0.59* (0.02)	0.59* (0.01)	0.98		
<b>yields</b>									
two years	-0.21 (0.47)	0.46* (0.12)	0.21	0.15 (0.20)	0.46* (0.05)	0.38* (0.02)	0.86	-0.12 (0.21)	0.83* (0.03)
five years	-0.27 (0.51)	0.26 (0.15)	0.07	0.11 (0.25)	0.26* (0.08)	0.40* (0.03)	0.76	-0.06 (0.28)	0.83* (0.04)
ten years	-0.33 (0.46)	0.13 (0.12)	0.02	-0.03 (0.29)	0.13 (0.07)	0.32* (0.03)	0.60	-0.14 (0.31)	0.70* (0.06)
<b>forward rates (6mo)</b>									
two years	-0.18 (0.60)	0.30 (0.16)	0.06	0.27 (0.28)	0.30* (0.07)	0.48* (0.03)	0.80	0.08 (0.29)	0.83* (0.04)
five years	-0.36 (0.49)	0.07 (0.14)	< 0.01	-0.03 (0.34)	0.07 (0.10)	0.35* (0.05)	0.50	-0.12 (0.39)	0.77* (0.06)
ten years	-0.45 (0.47)	-0.04 (0.07)	< 0.01	-0.28 (0.42)	-0.04 (0.06)	0.18* (0.04)	0.20	-0.24 (0.42)	0.40* (0.08)

The number of observations is 148. Numbers in parentheses are White standard errors. \* denotes statistical significance at the 1% level. Forward rates are for six-month loans ending at the specified maturities. For details see text.



**Table 1.3:** Effects of macroeconomic announcements

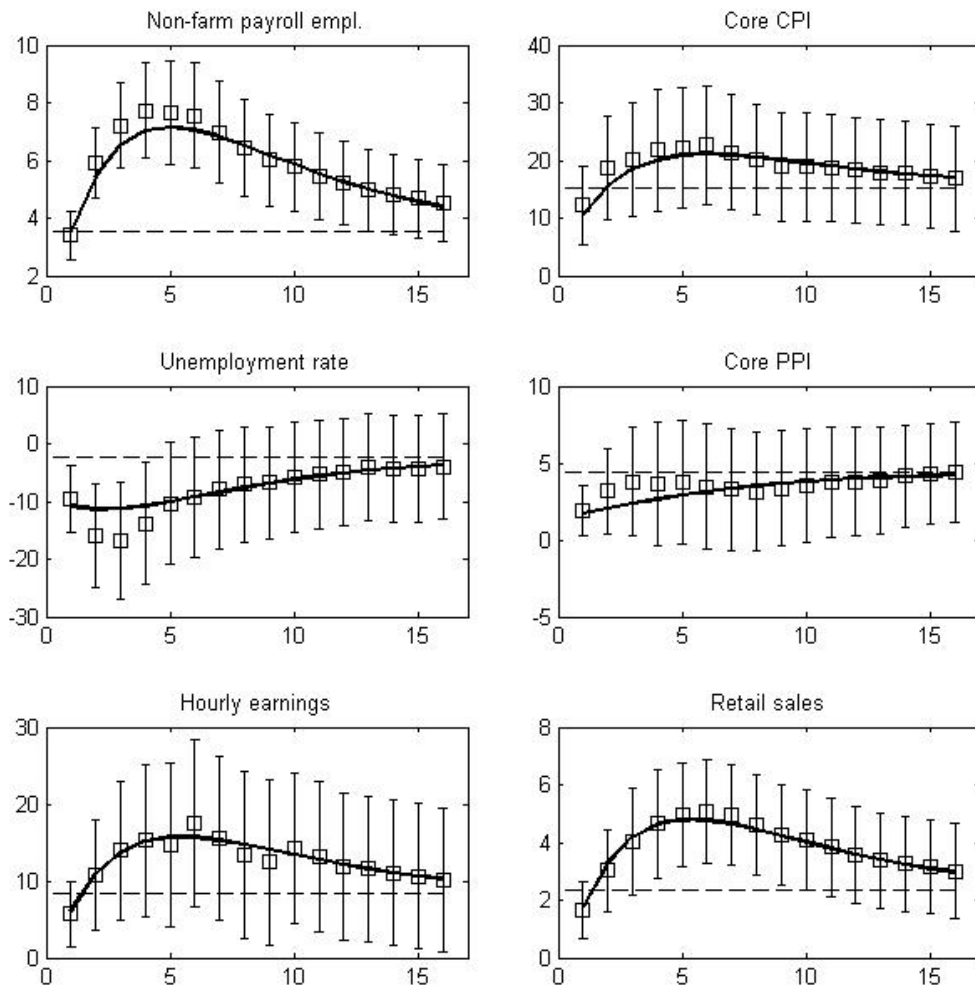
News release	FF2	ED4	2y yield	$\hat{\epsilon}_{1t}$	$\hat{\epsilon}_{2t}$	$\hat{\epsilon}_{3t}$
Non-farm payroll employment	3.09** (0.35)	8.11** (0.88)	5.82** (0.59)	3.76** (0.67)	-3.48** (1.00)	0.04** (0.01)
Unemployment rate	-1.30** (0.33)	-1.91* (0.74)	-1.44** (0.51)	-0.33 (0.62)	-0.89 (0.80)	-0.01 (0.00)
Hourly earnings	0.85** (0.26)	2.42** (0.81)	1.86** (0.58)	1.34 (0.73)	-1.64 (0.95)	0.01** (0.00)
Core CPI	1.02** (0.23)	2.58** (0.63)	2.08** (0.48)	1.83** (0.52)	-1.51** (0.57)	0.01** (0.00)
Core PPI	0.45* (0.19)	1.06 (0.59)	0.54 (0.48)	1.27** (0.47)	-0.86 (0.56)	-0.00 (0.00)
Retail sales	0.84** (0.23)	3.14** (0.65)	2.38** (0.50)	1.62** (0.55)	-2.14** (0.62)	0.02** (0.00)

Responses of instrument rates and term structure shocks to a one-standard-deviation surprise in six different macroeconomic data releases. Sample: Days with at least one macroeconomic announcement but without policy action, Oct-1988 to Jun-2007,  $N = 647$ . Numbers in parentheses are White standard errors. \* and \*\* denote significance at 5% and 1% level, respectively.

**Table 1.4:** Summary statistics for ratio of near-term rate changes

News	$\Delta ED_t^{(2)}/\Delta ED_t^{(1)}$			$\Delta ED_t^{(3)}/\Delta ED_t^{(1)}$			$\Delta ED_t^{(4)}/\Delta ED_t^{(1)}$			obs.
	med.	mean	(s.e.)	med.	mean	(s.e.)	med.	mean	(s.e.)	
Empl. report	1.83	2.30	(0.20)	2.25	3.08	(0.30)	2.32	3.35	(0.38)	136
CPI/PPI	1.63	1.98	(0.16)	2.00	2.74	(0.27)	2.33	3.30	(0.33)	185
Retail sales	2.11	2.46	(0.22)	3.33	3.80	(0.45)	3.75	4.33	(0.57)	61
Pooled	1.78	2.17	(0.11)	2.26	3.03	(0.19)	2.50	3.48	(0.23)	382

Medians, means, and standard errors for the means for relative rate changes. Sample: 10-1988 to 06-2007. I include only observations for which the denominator rate change is non-zero (using unity for the ratio when the denominator is zero still results in medians and means far above one).



**Figure 1.7:** Term structure of announcement effects

Responses to a one-standard-deviation surprise in six different macroeconomic data releases: Empirical responses of futures rates with 95% confidence intervals (error-bars) and model-implied responses of futures rates (solid lines). Units are basis points.

# Chapter 2

## Term Premia and the News

### Abstract

How do short rate expectations and term premia respond to news? Dynamic term structure models typically imply that the term premium accounts for most of the procyclical response of long-term interest rates, which is at odds with the conventional wisdom about bond risk premia. Bias and lack of precision in the estimated short rate dynamics make it difficult to interpret this evidence. This paper solves these problems by imposing plausible zero restrictions on the market prices of risk. The no-arbitrage assumption becomes useful for estimation, because information in the cross section helps to pin down the dynamics of the short rate. Inference about term premia is performed in a Bayesian framework based on Markov Chain Monte Carlo methods. This allows the researcher to select plausible restrictions and to correctly quantify statistical uncertainty. The main empirical result is that under the restrictions favored by the data the expectations component, and not the term premium, accounts for the majority of high-frequency movements of long rates and for essentially all of their procyclical response to macroeconomic news.

## 2.1 Introduction

Policy makers and academic researchers have keen interest in the estimation of term premia in long-term interest rates. Different approaches have been used for this purpose, such as return regressions,<sup>1</sup> no-arbitrage dynamic term structure models<sup>2</sup> (DTSMs) and more recently macro-finance term structure models.<sup>3</sup> Previous studies established that there is a sizeable term premium which varies over time. Thus the expectations hypothesis, which posits a constant term premium, is at odds with the data. Moreover the term premium seems to vary at business-cycle frequencies and this variation is countercyclical (Cochrane and Piazzesi, 2005; Piazzesi and Swanson, 2008).

An important question in this context is how the response of the term structure of interest rates to news events, such as macroeconomic announcements and policy actions, decomposes into revisions of short rate expectations and changes in term premia. More generally, how do term premia change at high frequencies? This question is of relevance both to policy makers, who need up-to-date information about market participants' expectations of the future policy path, and to investors, since optimal asset allocation requires knowing how new information changes expected returns. A recent attempt to answer this question was made by Meredith Beechey (2007), who uses the DTSM of Kim and Wright (2005) to decompose rate changes into expectations and term premium components. Her study finds that the procyclical response of long-term interest rates is mainly due to changes in term premia – short rate expectations seem to hardly react to the news. However, this strong procyclical response of term premia at high frequencies is a puzzle in light of the conventional term premium wisdom cited above.

This puzzling evidence results from two general problems with the estimation of DTSMs and term premia that are caused by the high persistence<sup>4</sup> of the

---

<sup>1</sup>See for example Fama and Bliss (1987), Cochrane and Piazzesi (2005) and Piazzesi and Swanson (2008).

<sup>2</sup>The standard reference for affine no-arbitrage term structure models is Duffie and Kan (1996), applications to term premium estimation include Dai and Singleton (2002) and Kim and Wright (2005).

<sup>3</sup>Studies that incorporate macro-factors into no-arbitrage models and study term-premium properties include Ang and Piazzesi (2003); Rudebusch et al. (2006); Joslin et al. (2008).

<sup>4</sup>The null of a unit root can usually not be rejected for the short rate, see for example Rose

short rate: *lack of precision* and *bias* in estimates of the short rate's dynamic properties. At high frequencies the persistence is particularly strong, so the resulting problems are even more severe. The literature has documented these issues, for example in Duffee and Stanton (2004) and Kim and Orphanides (2005), but so far has offered no solution. The present paper, which focuses on high-frequency changes in interest rates and risk premia, solves both of these problems.

Persistence leads to imprecise estimates of the unconditional mean and the speed of mean reversion of the short rate intuitively because the short rate does not revert to its mean very often. Term premium estimates rely on forecasts of the short rate, thus the high estimation uncertainty about short rate dynamics translates into equally large uncertainty about the term premium. Empirical results that proceed by using term premium estimates without accounting for their uncertainty (e.g. Beechey, 2007) should therefore be taken with a grain of salt. Unfortunately, as John Cochrane (2007, p. 278) puts it,

we are usually treated only to one estimate based on one a priori specification, usually in levels, and usually with no measure of the huge sampling uncertainty.

An indication of the magnitude of estimation and specification uncertainty are the strikingly different term premium estimates in the literature (Swanson, 2007). I will quantify this statistical uncertainty and show its relation to the pricing of risk.

A closely related problem is an upward bias in the estimated speed of mean reversion of the short rate: The largest root of a persistent variable is generally under-estimated. This results in short rate forecasts that revert to the unconditional mean too quickly. As a consequence, changes in term premia are usually found to be the dominant source of the variation in long-term interest rates (e.g. Kim & Wright, 2005). This is implausible since we think that the term premium moves slowly.

Estimation of a DTSM requires inference about the short rate dynamics under both the risk-neutral ( $\mathbb{Q}$ ) measure and the physical ( $\mathbb{P}$ ) measure. The  $\mathbb{Q}$ -dynamics determine the loadings of the cross section of interest rates on the term 

---

 (1988) and Jardet et al. (2009). Whereas the short rate is not literally  $I(1)$ , since it is bounded from below and usually remains in a certain range, its largest root is certainly very close to one.

structure factors and can be pinned down very precisely. The  $\mathbb{P}$ -dynamics describe the evolution of the factors over time and estimation is difficult for the above-mentioned reasons. Can we use the information in the cross section to improve our estimates of the  $\mathbb{P}$ -dynamics? No-arbitrage requires consistency between cross-sectional and dynamic properties of the term structure, allowing for a risk adjustment. If market prices of risk are unrestricted, then the parameters under the  $\mathbb{Q}$  measure and under the  $\mathbb{P}$  measures are estimated independently of each other – no-arbitrage has no bite in this case. The solution is to impose restrictions on the prices of risk. Whereas this fact has been recognized in the literature (Kim and Orphanides, 2007; Cochrane and Piazzesi, 2008), the important question about which restrictions we should impose has so far not been satisfactorily answered. A commonly employed approach is to restrict some risk sensitivity parameters with large standard errors to zero and then to re-estimate the model. This ad-hoc procedure, unsatisfactory for several reasons, usually leads to very few restrictions and does not solve the bias and uncertainty problems.<sup>5</sup>

This paper develops a new statistical framework for the estimation of DTSMs which allows the researcher to impose sensible restrictions on the prices of risk. The question posed by Cochrane (2007, p. 276), “Can statistics help us?”, is answered with a strong affirmative. The physical dynamics of the short rate are estimated more accurately and more precisely. The resulting decompositions of rate changes are more reliable, and they also turn out to be more plausible than those of conventional DTSMs. My approach opens up a new road for estimation of term structure models, where parsimony and no-arbitrage play a more prominent role than they currently do in the term structure literature.

Estimation and inference is performed using Markov Chain Monte Carlo (MCMC) methods. A Bayesian framework is necessary not only for correctly quantifying statistical uncertainty but more importantly for appropriately restricting the market prices of risk: The tools of Bayesian model selection allow me to select those zero restrictions on the risk sensitivity parameters that are most strongly sup-

---

<sup>5</sup>Prominent studies that employ this two-step approach are Dai and Singleton (2002); Ang and Piazzesi (2003); Kim and Wright (2005). None of these restrict more than three risk sensitivity parameters to zero.

ported by the data. I develop a new algorithm related to Gibbs Variable Selection, using latent indicator variables that represent the restrictions, in order to identify plausible models. For this smaller subset of candidate models I can then precisely estimate posterior model probabilities and assess their economic implications.

The DTSM used in this paper belongs to the affine Gaussian class. Its key distinguishing characteristic is that the  $\mathbb{Q}$ -dynamics are specified in a way that identifies the latent factors a priori as level, slope and curvature. This is crucial for two reasons: First, the presence of a level factor, an empirical necessity, requires a unit root for the short rate under  $\mathbb{Q}$ . Because the model is parameterized in terms of  $\mathbb{Q}$ -dynamics and prices of risk, it then depends on the restrictions on the prices of risk whether there is a unit root under  $\mathbb{P}$ . Hence I let the data choose between a stationary or integrated specification for the short rate, instead of imposing this a priori. Second, having labeled factors allows for an economic interpretation of risk premia, based on the inference about, for example, the importance of level vs. slope risk. The DTSM in this paper parallels the Dynamic Nelson-Siegel model in Christensen et al. (2007), but gives rise to new Nelson-Siegel-type factor loadings since it is set in discrete time.

I also differ from most previous studies in basing the analysis on Eurodollar futures, which have several practical advantages over Treasury bonds in this context: First the futures rates directly reflect the forward rate curve, whereas forward rates derived from bond prices depend on the algorithm used to infer zero rates from observed bond prices (smoothed vs. unsmoothed, etc.). Second, the liquidity is very high, in fact Eurodollar futures are the most liquid futures contracts worldwide (in terms of open interest). Third, the most liquidly traded government bonds, on-the-run Treasury securities, do not cover the maturity spectrum at a similar detail. Last, the futures contracts are not affected by flight-to-quality effects or other extraordinary forces affecting supply and demand of Treasury securities. A conceptual advantage is that the payoffs of money market futures depend linearly on future short rates, thus convexity terms, which necessarily arise for yields and forward rates implied by bond prices, are absent in the pricing formulas

for these securities.<sup>6</sup>

I decompose rate changes into three components: Revisions of short rate expectations, surprise changes in the term premium and expected returns. The usual decomposition into predictable and unpredictable components overlooks that unpredictable changes have two different sources. Since the predictable component is negligible at high frequencies, the question is how much changes in short rate expectations vs. surprise term premium changes contribute to observed rate changes. Because of the bias problem unrestricted models wrongly attribute the majority of long-maturity rate changes to the term premium.

The paper also introduces some new ways to represent the implications of different models visually. I derive a “risk-neutral volatility curve”, which captures the volatility of short rate expectations across maturities. Based on the analytical decompositions I graphically summarize the contribution of short rate expectations to rate changes on individual days, to the level of volatility, and to the systematic response of forward rates to macroeconomic news.

Turning to the empirical findings of the paper, a key result is that *the data support strongly restricted prices of risk*. This is plausible if we believe in no-arbitrage: physical and risk-neutral dynamics *should* be close to each other. As a consequence of the restrictions the physical dynamics are estimated with higher precision, and our inference about the term premium becomes more reliable. This is the key contribution of the paper: by imposing sensible zero restrictions on risk sensitivity parameters we can overcome the bias and uncertainty problems that most DTSMs suffer from. Those restrictions that imply a unit root for the short rate receive particularly strong empirical backing – a stochastically trending short rate evidently is a good approximation to the true data-generating process, at least at the daily frequency.

Changes in short rate expectations account for most of the daily volatility in forward rates across all maturities. Thus *the procyclical response of long forward rates to macro news is mainly due to revisions of short rate expectations* – I overturn the result that this response primarily reflects changes in the term premium.

---

<sup>6</sup>Hence Nelson-Siegel loadings without any convexity adjustments are consistent with no-arbitrage.



The contribution of risk premia to variation in long rates is found to be much smaller than implied by most DTSMs in the literature. My results accord with our intuition in two ways: First, we think that term premia move in a countercyclical fashion, thus their contribution to procyclical interest rate changes caused by news should be small. Second, since most macroeconomists would agree that term premia probably move at business cycle frequencies, we would not expect them to account for much variability at the daily frequency. The decompositions of daily forward rate changes into risk premium and expectations components proposed in this paper are thus more plausible than those implied by the largely unrestricted DTSMs that are common in the literature.

These results remain robust when I account for specification uncertainty. The decompositions obtained using Bayesian model averaging (BMA) confirm an important role of short rate expectations for variation in long rates. Importantly the use of BMA solves the “discontinuity problem”, the stark difference between forecasts for the short rate depending on whether its largest root is less than or equal to one (Cochrane and Piazzesi, 2008; Jardet et al., 2009), since estimates effectively are averages of stationary and non-stationary specifications.

The paper is structured as follows: Section 2.2 describes the DTSM and shows how it can be used to decompose observed rate changes and the term structure of volatility. In Section 2.3 the model is estimated without restrictions on the prices of risk, revealing dramatic uncertainty and bias in decompositions of rate changes. Section 2.4 develops and applies a new framework for estimation of DTSMs, based on restrictions on the prices of risk. Section 2.5 concludes.

## 2.2 Dynamic Term Structure Model

In this section I introduce the affine Gaussian DTSM with its particular specification of the risk-neutral dynamics, present a new decomposition of changes in forward rates, introduce the “risk-neutral vol curve” and discuss the pricing of Eurodollar futures.

### 2.2.1 Affine Gaussian DTSMs

Denote by  $X_t$  the  $(k \times 1)$  vector of term structure factors which represent the new information that market participants obtain at time  $t$ . Generally it can contain both latent and observable factors, but this paper uses only latent factors. Assume that  $X_t$  follows a first-order Gaussian vector autoregression under the physical measure  $\mathbb{P}$ :

$$X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t, \quad (2.1)$$

with  $\varepsilon_t \sim N(0, I_k)$  and  $E(\varepsilon_r \varepsilon_s') = 0$ ,  $r \neq s$ . The frequency of the model is daily. The short rate  $r_t$ , the overnight rate, is specified to be an affine function of the factors:

$$r_t = \delta_0 + \delta_1' X_t. \quad (2.2)$$

This short rate is the policy instrument of the Federal Reserve, thus expectations under the physical measure of its future values correspond to expectations about future monetary policy.<sup>7</sup>

Assuming absence of arbitrage there exists a risk-neutral probability measure, denoted by  $\mathbb{Q}$ , which prices all financial assets. Equivalently, there is a stochastic discount factor (SDF) that defines the change of probability measure between the physical and the risk-neutral world. The one-period SDF,  $M_{t+1}$ , is specified as

$$-\log M_{t+1} = r_t + \frac{1}{2} \lambda_t' \lambda_t + \lambda_t' \varepsilon_{t+1}, \quad (2.3)$$

with the  $(k \times 1)$  vector  $\lambda_t$ , the prices of risk, being an affine function of the factors,

$$\lambda_t = \lambda_0 + \lambda_1 X_t. \quad (2.4)$$

The risk sensitivity parameters  $\lambda_0$  ( $k \times 1$ ) and  $\lambda_1$  ( $k \times k$ ) determine the behavior of risk premia. Under these assumptions the risk-neutral dynamics (see Appendix 2.6.1) are given by

$$X_t = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_{t-1} + \Sigma \varepsilon_t^{\mathbb{Q}}, \quad (2.5)$$

---

<sup>7</sup>I abstract from the facts that the overnight rate in the U.S., the effective fed funds rate, deviates from the target set by the monetary authority, and that the target has a step-function character. Both simplifications are inconsequential since I do not include observations of the short rate – inference is based on Eurodollar futures rates, which correspond to average forward rates over an entire quarter.

where  $\varepsilon_t^{\mathbb{Q}} \overset{\mathbb{Q}}{\sim} N(0, I_k)$ ,  $E^{\mathbb{Q}}(\varepsilon_r^{\mathbb{Q}} \varepsilon_s^{\mathbb{Q}'}) = 0$ ,  $r \neq s$ , and the parameters describing the physical and risk-neutral dynamics are related in the following way:

$$\mu^{\mathbb{Q}} = \mu - \Sigma \lambda_0, \quad \Phi^{\mathbb{Q}} = \Phi - \Sigma \lambda_1. \quad (2.6)$$

This discrete-time affine Gaussian DTSM was introduced by Ang and Piazzesi (2003) and is now widely used. While it does not allow for stochastic volatility it is rather flexible in matching risk premia (Dai and Singleton, 2003).<sup>8</sup>

No-arbitrage requires the consistency of dynamic properties (determined by  $\mu$  and  $\Phi$ ) and cross-sectional properties (determined by  $\mu^{\mathbb{Q}}$  and  $\Phi^{\mathbb{Q}}$ ) of interest rates, allowing for a risk-adjustment, and (2.6) makes this risk-adjustment explicit. Note however that if  $\lambda_0$  and  $\lambda_1$  are left unrestricted, then (2.6) is not restrictive at all: Any estimates for the physical dynamics are consistent, for some prices of risk, with a given choice of risk-neutral dynamics. Most studies impose no or only minimal restrictions on  $\lambda_0$  and  $\lambda_1$ . This paper will show that strong zero-restrictions on  $\lambda_1$  are supported by the data, and that imposing these restrictions provides us with more precise and more plausible term premium estimates.

## 2.2.2 Forward rates

With this framework we can price any asset with payoff depending on the future path of the short rate, the price being given by the discounted expected future payoff under  $\mathbb{Q}$ . Instead of considering bonds, this paper focuses on money market futures, which pay off according to the average short rate over a future time horizon. Thus the main object of interest is the expected future short rate under the risk-neutral measure,  $f_t^n = E_t^{\mathbb{Q}}(r_{t+n})$ . I will refer to this object as a *forward rate*, although by this term one would usually mean the rate that can be contracted today for a loan from  $t + n$  to  $t + n + 1$  by entering the appropriate bond positions and which includes Jensen inequality terms.<sup>9</sup>

<sup>8</sup>Examples of studies that assess the behavior of the term premium using this affine Gaussian framework are Duffee (2002); Ang and Piazzesi (2003); Kim and Orphanides (2005); Kim and Wright (2005); Rudebusch and Wu (2008); Joslin et al. (2008).

<sup>9</sup>The actual forward rate based on bond positions is equal to  $\log \frac{P_t^n}{P_t^{n+1}} = \log \frac{E_t^{\mathbb{Q}} \exp(-r_t - r_{t+1} - \dots - r_{t+n-1})}{E_t^{\mathbb{Q}} \exp(-r_t - r_{t+1} - \dots - r_{t+n})}$  where  $P_t^n$  is the time- $t$  price of a discount bond with  $n$  days until

Solving for these forward rates is straightforward because of the linearity of the short rate and the availability of analytical expressions for the conditional expectation of  $X_t$ . We obtain

$$\begin{aligned} f_t^n &= E_t^{\mathbb{Q}}(\delta_0 + \delta'_1 X_{t+n}) \\ &= \delta_0 + \delta'_1 \left[ \sum_{i=0}^{n-1} (\Phi^{\mathbb{Q}})^i \mu^{\mathbb{Q}} + (\Phi^{\mathbb{Q}})^n X_t \right] \\ &= A_n + B'_n X_t \end{aligned} \tag{2.7}$$

$$A_n = \delta_0 + \delta'_1 \left[ \sum_{i=0}^{n-1} (\Phi^{\mathbb{Q}})^i \mu^{\mathbb{Q}} \right], \quad B'_n = \delta'_1 (\Phi^{\mathbb{Q}})^n.$$

These loadings correspond to those for one-period forward rates common in the bond pricing literature, as derived for example in Cochrane and Piazzesi (2008), with the difference that Jensen inequality terms resulting from convexity effects are absent in our case.

### 2.2.3 The arbitrage-free Dynamic Nelson-Siegel model

Since not all parameters of the DTSM are identified, some normalization restrictions need to be imposed (Dai and Singleton, 2000). A DTSM is called “canonical” or “maximally flexible” if there are no over-identifying restrictions, which for the case  $k = 3$  amounts to 22 free parameters in  $(\delta_0, \delta_1, \mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}}, \mu, \Phi, \Sigma)$ . A particularly convenient normalization is the canonical form of Joslin et al. (2009) (JSZ), who impose the restrictions  $\mu^{\mathbb{Q}} = 0, \delta_1 = \iota_k$  and parameterize the  $\mathbb{Q}$ -dynamics in terms of  $\delta_0$  and the eigenvalues of  $K^{\mathbb{Q}} = \Phi^{\mathbb{Q}} - I_k$ , which is taken to be in real ordered Jordan form.<sup>10</sup>

In canonical DTSMs with only latent factors the role of each factor is a priori left unidentified. On the other hand the widely used yield-curve parametrization of Nelson and Siegel (1987) posits a simple factor structure for forward rates involving three factors that correspond to level, slope and curvature. The dynamic version

---

maturity.

<sup>10</sup>All normalizations are imposed on the  $\mathbb{Q}$ -dynamics, hence for given (observable or filtered) factors and absent restrictions on the prices of risk, consistent estimates of  $(\mu, \Phi)$  can be obtained using ordinary least squares, as shown by JSZ.

of the original Nelson-Siegel forward rate curve,

$$f_t^n = X_t^{(1)} + e^{-\lambda n} X_t^{(2)} + \lambda n e^{-\lambda n} X_t^{(3)},$$

is implied by a continuous-time three-factor affine DTSM with a specific choice for the risk-neutral dynamics, as shown by Christensen et al. (2007). My discrete-time analogue to their model, using a discretization scheme with one-period increments, is given by the following specification:

$$\delta_0 = 0, \delta_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mu^{\mathbb{Q}} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Phi^{\mathbb{Q}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho & 1 - \rho \\ 0 & 0 & \rho \end{pmatrix} \quad (2.8)$$

where the parameter  $\rho$  is restricted to be less than one in absolute value. This is the Dynamic Nelson-Siegel (DNS) specification in discrete time. Straightforward algebra based on equation (2.7) leads to a Nelson-Siegel-type forward rate curve given by

$$f_t^n = X_t^{(1)} + \rho^n X_t^{(2)} + n(1 - \rho)\rho^{n-1} X_t^{(3)}. \quad (2.9)$$

Notably there is no convexity term since I consider  $f_t^n = E_t^{\mathbb{Q}}(r_{t+n})$ .

The DNS specification in (2.8) is equivalent to the JSZ normalization with three over-identifying restrictions (JSZ, Section 4.2). The first restriction is the unit eigenvalue of  $\Phi^{\mathbb{Q}}$ , i.e. a zero eigenvalue of  $K^{\mathbb{Q}}$ . The empirical evidence (e.g. Gürkaynak et al., 2005b) overwhelmingly calls for a unit root under  $\mathbb{Q}$  since otherwise long-horizon forward rates would be constant. The yield curve cannot have a level factor unless  $\Phi^{\mathbb{Q}}$  has a unit eigenvalue. The second restriction is the zero long-run mean of the short rate under  $\mathbb{Q}$  ( $\delta_0 = 0$ ). Given that  $X_t^{(1)}$  serves as a level factor there is no need for a non-zero unconditional mean. The unit root and zero mean under  $\mathbb{Q}$  are highly plausible and empirically necessary restrictions on the JSZ normalization.

The third restriction of the DNS model is that the two other eigenvalues of  $\Phi^{\mathbb{Q}}$  are equal, which identifies the third factor as curvature. This somewhat restrictive assumption is useful because now the factors are a priori identified as level, slope and curvature. One advantage of this is that it enables us to provide

an economic interpretation of risk premia.<sup>11</sup> Labeling the factors by means of a particular choice of the risk-neutral dynamics, as well as introducing a unit root under  $\mathbb{Q}$ , are the two major advantages of the DNS specification.

An important remark about the consequence of a unit root under  $\mathbb{Q}$ : Because I will estimate the model based on a parametrization in terms of  $(\lambda_0, \lambda_1)$  and not in terms of  $(\mu, \Phi)$ , and since  $\Phi = \Phi^{\mathbb{Q}} + \Sigma\lambda_1$ , it will depend entirely on the restrictions imposed on  $\lambda_1$  whether  $\Phi$  has a unit eigenvalue. Absent any restrictions, estimates will generally imply a stationary short rate. By having the data choose plausible zero restrictions on  $\lambda_1$  the possibility of a unit root under  $\mathbb{P}$  will be explicitly allowed for.

#### 2.2.4 Risk-neutral rates and forward risk premia

The term premium can be defined in different ways, and I focus on the forward risk premium, which I relate later to the return risk premium.<sup>12</sup> If the marginal investor was risk-neutral, the forward rate  $f_t^n$  would be equal to the expected future short rate *under the physical measure*. This is usually called the “risk-neutral forward rate”, here denoted by  $\tilde{f}_t^n$ . We have

$$\tilde{f}_t^n = E_t(r_{t+n}) = \delta_0 + \delta_1' E_t(X_{t+n}) = \tilde{A}_n + \tilde{B}_n' X_t,$$

with loadings given by

$$\tilde{A}_n = \delta_0 + \delta_1' \left[ \sum_{i=0}^{n-1} \Phi^i \mu \right], \quad \tilde{B}_n' = \delta_1' \Phi^n.$$

Risk-neutral rates embody the expectations about future short rates and thus about future monetary policy. They are not observable and have to be inferred by constructing forecasts for the short rate. The forward risk premium, denoted

---

<sup>11</sup>Specifically we would like to know whether level risk or slope risk causes time-variation in the term premium, since this provides some hints about the role of different macroeconomic shocks, as pointed out by Cochrane and Piazzesi (2008).

<sup>12</sup>Specifically, the term premium can be defined equivalently as a yield risk premium (the difference between a bond yield and the average expected future short rate), a forward risk premium (the difference between a forward rate and the expected future short rate), or a return risk premium (the expected excess return on a bond or futures contract). For a detailed discussion of how these are related see Cochrane and Piazzesi (2008).

by  $\Pi_t^n$ , is defined as the difference between the forward rate and the risk-neutral forward rate:

$$\Pi_t^n = f_t^n - \tilde{f}_t^n = A_n - \tilde{A}_n + (B_n - \tilde{B}_n)' X_t.$$

Importantly, the statistical uncertainty about  $\mu$  and  $\Phi$  translates into uncertainty about risk-neutral rates and forward risk premia.

## 2.2.5 Decomposing rate changes

The main question asked in this paper is to what extent daily changes in the forward rate curve are driven by changing short rate expectations. By definition, rate changes decompose into changes in risk-neutral rates and changes in forward risk premia,  $f_{t+1}^n - f_t^{n+1} = \tilde{f}_{t+1}^n - \tilde{f}_t^{n+1} + \Pi_{t+1}^n - \Pi_t^{n+1}$ . The DTSM can be used to provide a more detailed decomposition:

$$\begin{aligned} f_{t+1}^n - f_t^{n+1} &= A_n + B_n' X_{t+1} - A_{n+1} - B_{n+1}' X_t \\ &= B_n' \Sigma \varepsilon_{t+1}^Q \\ &= \tilde{B}_n' \Sigma \varepsilon_{t+1} + (B_n - \tilde{B}_n)' \Sigma \varepsilon_{t+1} + B_n' \Sigma \lambda_t. \end{aligned} \quad (2.10)$$

This rate change corresponds to the one-period return of a hypothetical futures contract which pays the difference between the realized future short rate and the contracted rate.<sup>13</sup> Notably it is an *excess return*, because the risk-neutral expected return,  $E_t^Q(f_{t+1}^n - f_t^{n+1})$  is zero. Expression (2.10) decomposes this return into three components:

*Revisions to short rate expectations* – The first component corresponds to the change in the expectation of the short rate for  $t + n + 1$ :

$$\begin{aligned} (E_{t+1} - E_t) r_{t+n+1} &= \delta_1' (E_{t+1} X_{t+n+1} - E_t X_{t+n+1}) \\ &= \delta_1' \Phi^n \Sigma \varepsilon_{t+1} = \tilde{B}_n' \Sigma \varepsilon_{t+1}. \end{aligned}$$

This component, which equals the change in the risk-neutral rate  $\tilde{f}_{t+1}^n - \tilde{f}_t^{n+1}$ ,

---

<sup>13</sup>Specifically it is the absolute return if one enters at  $t$  into a long position in a contract that pays  $r_{t+n+1} - f_t^{n+1}$  at maturity, and liquidates the position at  $t + 1$ . Note that money market futures usually pay the difference between contracted rate and short rate, in which case the above rate change corresponds to the return on a short position.

captures how market participants revise their expectations of future monetary policy.

*Surprise changes in the forward risk premium* – The second component is equal to the unexpected change in the forward risk premium:

$$\begin{aligned} (\Pi_{t+1}^n - E_t \Pi_{t+1}^n) &= (B_n - \tilde{B}_n)' (X_{t+1} - E_t X_{t+1}) \\ &= (B_n - \tilde{B}_n)' \Sigma \varepsilon_{t+1}. \end{aligned}$$

*Expected returns* – The third component is equal to the expected change in the forward risk premium:

$$\begin{aligned} E_t \Pi_{t+1}^n - \Pi_t^{n+1} &= E_t f_{t+1}^n - f_t^{n+1} = A_n + B_n' E_t X_{t+1} - A_{n+1} - B_{n+1}' X_t \\ &= B_n' E_t \varepsilon_{t+1}^Q = B_n' \Sigma \lambda_t = B_n' \Sigma (\lambda_0 + \lambda_1 X_t) \end{aligned}$$

This term captures the predictable part of the daily return and corresponds to the *return risk premium*.<sup>14</sup> For daily rate changes this component is negligibly small. How large it is for longer holding periods is an important question which I consider in Section 2.4.6.

Using equation (2.10) changes in forward rates can be decomposed into expectations and risk premium components. Importantly this decomposition will inherit the statistical uncertainty from the inference about the unknown parameters and factors. I will show below the dramatic uncertainty in DTSMs that lack the appropriate restrictions on the prices of risk.

To foreshadow a key empirical result: Typically DTSMs imply that for long maturities the first component is small and the second component accounts for most of daily rate changes. This is puzzling given the conventional wisdom about bond risk premia. If restrictions are imposed on the prices of risk a much larger share of rate changes is attributed to the first component, changes in expectations, implying little daily variability of the term premium.

---

<sup>14</sup>In the language of Cochrane and Piazzesi (2005),  $B_n' \Sigma \lambda_1 X_t$  is the *return-forecasting factor*, which generally differs across maturities. It differs across maturities only by a factor of proportionality if only one element in the vector  $\lambda_t$  is non-zero, see for example the model of Cochrane and Piazzesi (2008).



## 2.2.6 Term Structure of Volatility

The term structure of volatility, the “vol curve”, describes the volatility of changes in yields or forward rates across maturities, either in the sample or in population. Given the decomposition in (2.10) the variance of forward rate changes is given by

$$\text{Var}(f_{t+1}^n - f_t^{n+1}) = B_n' \Sigma (I_k + \text{Var}(\lambda_t)) \Sigma' B_n$$

The term structure of volatility is the square root of this expression for varying  $n$ . Variability of forward rates is driven both by an unpredictable component, the innovations to the factors, and by a predictable component, the variation in the prices of risk. It will turn out that the predictability of daily changes is very small, thus  $\text{Var}(f_{t+1}^n - f_t^{n+1}) \approx B_n' \Sigma \Sigma' B_n$ .

Our framework allows us to assess the importance of changes in short rate expectations and forward risk premia for variability in forward rates. Specifically, we can calculate the term structure of volatility that would prevail if forward rates were only driven by changes in short rate expectations, i.e. if term premia were constant. The variance of changes in risk-neutral forward rates is

$$\text{Var}(\tilde{f}_{t+1}^n - \tilde{f}_t^{n+1}) = \text{Var}(\tilde{B}_n' \Sigma \varepsilon_{t+1}) = \tilde{B}_n' \Sigma \Sigma' \tilde{B}_n$$

and I will call the square root of this expression for varying  $n$  the “risk-neutral vol curve”.

## 2.2.7 Eurodollar futures

In order to estimate the term structure model and for all subsequent empirical analysis, this paper uses Eurodollar futures contracts.<sup>15</sup> These instruments settle based on the 3-month LIBOR rate at some future date (the settlement day). This settlement rate can safely be taken to be the average expected short rate (under  $\mathbb{Q}$ ) for the 3-month period following the settlement day:  $S_t = N^{-1} \sum_{h=0}^{N-1} E_t^{\mathbb{Q}}(r_{t+h})$ , where  $N$  is the number of days in the quarter, taken

<sup>15</sup>For detailed information on Eurodollar futures contracts please refer to the Chicago Mercantile Exchange’s web site at <http://www.cmegroup.com/trading/interest-rates/stir/eurodollar.html> (accessed 08/29/2009).

to be 91 throughout this paper.<sup>16</sup> Eurodollar futures contracts involve no cost today and have payoff proportional to the difference between contracted rate and settlement rate.<sup>17</sup> For the Eurodollar futures contract that settles at the end of quarter  $i$ , where  $i = 1$  corresponds to the current quarter, we have the following fundamental pricing equation:

$$0 = E_t^{\mathbb{Q}}(ED_t^{(i)} - S_{T(i,t)}),$$

where  $ED_t^{(i)}$  is the futures rate and  $T(i,t)$  denotes the settlement day that corresponds to contract  $i$  on day  $t$ . Settlement takes place on the last day of the quarter, therefore  $T(i,t) = t + iN - d(t)$ , where  $d(t)$  is the day within the quarter of calendar day  $t$ . The futures rate is thus given by

$$\begin{aligned} ED_t^{(i)} &= E_t^{\mathbb{Q}}(S_{T(i,t)}) \\ &= N^{-1} \sum_{n=iN-d(t)}^{(i+1)N-d(t)-1} E_t^{\mathbb{Q}} r_{t+n} \\ &= N^{-1} \sum_{n=iN-d(t)}^{(i+1)N-d(t)-1} (A_n + B'_n X_t) \\ &= a_i + h'_i X_t. \end{aligned} \tag{2.11}$$

Note that the futures rate is simply the average of the  $N$  relevant forward rates for the three-month period starting on the settlement day. The scalar  $a_i$  and the vector  $h_i$  are the averages of  $A_n$  and  $B_n$ , respectively, over the relevant period.<sup>18</sup>

If market participants were risk-neutral, the futures rates would be equal

---

<sup>16</sup>The LIBOR rate is usually very closely related to the average expected effective federal funds rate. The difference between the two, which is measured by the so-called LIBOR-OIS spread, stems from a small term premium and a credit risk premium due to the three-month commitment at a specific rate with a particular counter-party when lending at LIBOR. This spread was very small (around 8 basis points) and showed little variability throughout the period of our data set, which ends before the start of the recent financial turmoil.

<sup>17</sup>We abstract from the fact that in reality payments are made every day because of marking-to-market. Evidence in Piazzesi and Swanson (2008) indicates that this effect is likely to be negligible in our context.

<sup>18</sup>The last equality is an approximation due to the fact that instead of having different  $a_i$ 's and  $h_i$ 's depending on the day of the quarter, I set  $d(t)$  equal to the constant 45 (approximately the average of  $d(t)$ ), which leads to only a very small approximation error and significantly lowers the computational burden.

to expected average future short rates. This risk-neutral futures rate is given by

$$\tilde{ED}_t^{(i)} = \tilde{a}_i + \tilde{h}_i X_t \quad (2.12)$$

where  $\tilde{a}_i$  and  $\tilde{h}_i$  are the averages of  $\tilde{A}_n$  and  $\tilde{B}_n$ . The corresponding forward risk premium for contract  $i$  is given by  $\Pi_t^{(i)} = ED_t^{(i)} - \tilde{ED}_t^{(i)}$ . The decomposition for changes in forward rates in (2.10) analogously holds for changes in futures rates:

$$ED_{t+1}^{(i)} - ED_t^{(i)} = \tilde{h}_i' \Sigma \varepsilon_{t+1} + (h_i - \tilde{h}_i)' \Sigma \varepsilon_{t+1} + h_i' \Sigma \lambda_t. \quad (2.13)$$

The term structure of volatility and the risk-neutral vol curve for Eurodollar futures are analogous to those for forward rates, with  $B_n$  and  $\tilde{B}_n$  replaced by  $h_i$  and  $\tilde{h}_i$ . We now have the necessary theoretical foundations and can proceed by estimating the model.

## 2.3 Estimation of the unrestricted model

Turning to the estimation of the model, I first focus on a specification of the DTSM without any restrictions on the prices of risk. This serves as a benchmark and will reveal the bias and the large uncertainty underlying conventional term premium estimates.

The data set consists of daily observations on the rates for Eurodollar futures contracts maturing at the end of the current and the following 15 quarters, denoted by ED1 to ED16, thus covering the forward rate curve up to a maturity of about four years. The sample starts on 1 January 1990 and ends on 29 June 2007, before the start of the financial crisis. The number of days in the sample is  $T = 4401$ .

### 2.3.1 Econometric methodology

A state-space representation of the DTSM forms the basis for estimation. Equation (2.1) is the transition equation for the  $k \times 1$  state vector, which I reproduce here for convenience:

$$X_t = \mu + \Phi X_{t-1} + v_t, \quad v_t \sim N(0, Q), \quad E(v_t v_s') = 0, \quad t \neq s,$$

introducing the notation  $v_t = \Sigma \varepsilon_t$  and  $Q = \Sigma \Sigma'$ . The observation equation is

$$Y_t = a + H' X_t + w_t, \quad w_t \sim N(0, R), \quad E(w_t w_s') = 0, \quad t \neq s, \quad (2.14)$$

where  $Y_t$  is an  $m \times 1$  vector of observations at time  $t$ . We have  $m = 16$ , the observations being the rates of the 16 Eurodollar futures contracts,  $Y_t = (ED_t^{(1)}, \dots, ED_t^{(16)})'$ . The vector  $a$  stacks the intercepts  $a_1$  to  $a_{16}$ , and given the normalization  $\mu^{\mathbb{Q}} = 0$  we have  $a = 0$ . The  $k \times m$  coefficient matrix  $H = (h_1, \dots, h_{16})$  is determined by  $(\mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}})$ . The vector  $w_t$  contains measurement errors, included to avoid stochastic singularity as is common in the DTSM literature, which are serially uncorrelated and orthogonal to  $X_t$ . The variance-covariance matrix of  $w_t$ , denoted by  $R$ , is diagonal, and for the sake of parsimony I impose  $R = \sigma_w^2 I_m$ .

I parameterize the model in terms of the  $\mathbb{Q}$ -dynamics and the risk sensitivity parameters, which Bertholon et al. (2008) call the *back-modeling strategy*, for two reasons: First, the  $\mathbb{Q}$ -dynamics are chosen according to the DNS specification in order to a priori identify the factors as level, slope and curvature. Second, the focus of this paper is on inference and restrictions on the prices of risk, thus the parameters of the model need to explicitly include  $\lambda_0$  and  $\lambda_1$ . The parameters of our model, under the aforementioned normalization restrictions, are  $\theta = (\rho, \lambda_0, \lambda_1, \Sigma, \sigma_w^2)$ . The parameter  $\rho$  determines  $\Phi^{\mathbb{Q}}$ ,  $\mu^{\mathbb{Q}}=0$ , and together with the prices of risk and  $\Sigma$  the parameters of the physical dynamics are determined.

The model could in principle be estimated using classical statistical methods such as maximum likelihood (ML) estimation, but there are several shortcomings of this approach for estimation of DTSMs. The likelihood function generally has many dimensions (there are 20 parameters in our case) and is highly non-linear, which makes numerical optimization expensive and finding the global maximum difficult (Duffee and Stanton, 2004; Duffee, 2009). More specifically, the likelihood function of a DTSM oftentimes has “multiple inequivalent local maxima which have similar likelihood values but substantially different implications for economic quantities of interest” (Kim and Orphanides, 2005, p.10). Attempts to estimate the model with ML confirmed this: the likelihood function has several local maxima with different prices of risk, which confirms that the physical dynamics are hard

to pin down.<sup>19</sup> Joslin et al. (2009) have shown that for their canonical model the estimates of the P-dynamics are given by ordinary least squares, which simplifies estimation and can solve some of the above problems. However, their approach is based on the separation of risk-neutral and physical dynamics and thus is only applicable if the prices of risk are unrestricted.

Another important shortcoming of classical methods in the context of DTSM estimation is that one cannot correctly quantify the estimation uncertainty inherent in the calculations that use the model output. If for example we obtain ML estimates and calculate the risk-neutral volatility curve, which is a highly non-linear function of the model parameters, quantifying the estimation uncertainty requires approximation based on the delta-method, which would likely be unreliable. More importantly, quantifying the uncertainty that results from inference about both parameters and latent factors is not possible with classical methods, since latent factors are inferred conditional on given point estimates of the model parameters (Kim and Nelson, 1999, chap. 8).<sup>20</sup>

Bayesian estimation of the model by means of Markov Chain Monte Carlo (MCMC) methods overcomes these challenges.<sup>21</sup> It has numerous advantages: Computationally it is less challenging since it amounts to successively drawing parameters (from their conditional posterior distributions) instead of numerically maximizing a high-dimensional and strongly non-linear likelihood function. We can diagnose whether the MCMC algorithm is likely to have converged, whereas for MLE it is rather difficult to assess whether a solution is a global maximum or not. Instead of leading to multiple local maxima, the fact that the risk sensitivity parameters are hard to estimate is appropriately reflected in a rather flat posterior. The two biggest advantages of the MCMC approach in the present context are

---

<sup>19</sup>Christensen et al. (2007) claim that the a priori identification of the factors as level, slope and curvature solves this problem, however my own experience with the DNS model of this paper indicates that this is not the case – troublesome local maxima are still present.

<sup>20</sup>As an example consider the decomposition of rate changes according to equation (2.10): The uncertainty underlying the decomposition is due to uncertainty both in our estimates of latent factors (which determine  $\varepsilon_{t+1}$  and  $\varepsilon_{t+1}^Q$ ) and of the parameters (which determine  $B_n$ ,  $\tilde{B}_n$  and  $\Sigma$ ).

<sup>21</sup>For surveys on the use of MCMC methods in econometrics see Chib and Greenberg (1996) and Chib (2001). On estimation of asset pricing models using MCMC see Johannes and Polson (2009). Other authors that have used MCMC methods in the estimation of DTSMs are Ang et al. (2007) and Boivin et al. (2009).

the possibility to correctly account for estimation uncertainty, since the algorithm provides a sample from the joint posterior of the model parameters and the latent factors, and the fact that it will allow us to flexibly handle the issues of model choice and model uncertainty (Section 2.4).

The algorithm used to estimate the model is a block-wise Metropolis-Hastings (M-H) sampler, the details of which are described in Appendix 2.6.2. It provides us with a sample from the joint posterior distribution of model parameters and latent factors.

### 2.3.2 Parameter estimates

Table 2.1 presents the parameter estimates of the DTSM without any restrictions on the prices of risk. For each parameter I report the posterior mean and a 95% credibility interval (CI), obtained by taking the sample mean and appropriate quantiles from the MCMC sample.

The risk-neutral dynamics are estimated very precisely, as was to be expected: The loadings in  $H$  are determined by  $\rho$ , and the information in the cross-section of futures rates pins down this parameter very precisely.

The risk sensitivity parameters on the other hands are estimated very imprecisely. The CIs are large relative to the magnitudes of the estimates and for most parameters the CIs contain zero. The reason is the high persistence of the short rate: Intuitively, because the short rate does not revert to its mean very often it is hard to estimate its unconditional mean and its speed of mean reversion. Hence  $\mu$  and  $\Phi$  are estimated very imprecisely (see for example Duffee & Stanton, 2004). Since for given  $(\mu^Q, \Phi^Q)$  the  $\mathbb{P}$ -dynamics are determined by  $\lambda_0$  and  $\lambda_1$ , there is large estimation uncertainty about the risk sensitivity parameters.

Since the factors are a priori identified as level, slope and curvature, these estimates can help to understand the sources of time-variation in risk premia. The driving force seems to be changes in the prices of slope risk and curvature risk – only elements in the second and third row of  $\lambda_1$  are significantly different from zero. The price of level risk on the other hand does not seem to change over time. This contrasts with the results of Cochrane and Piazzesi (2008) who find that in

**Table 2.1:** Parameter estimates for unrestricted model

<i>Risk-neutral dynamics</i>				
$\rho$	.9973	[.9973, .9973]		
<i>Risk sensitivity parameters</i>				
	$\lambda_0$		$\lambda_1$	
	constant	level	slope	curvature
level risk	.1210	-.0164	.0103	.0014
	[-.03, .28]	[-.04, .01]	[-.01, .03]	[-.02, .02]
slope risk	<b>.3042</b>	<b>-.0444</b>	-.0190	<b>.0296</b>
	[.15, .46]	[-.07, -.02]	[-.04, .00]	[.01, .05]
curvature risk	.1273	<b>-.0239</b>	.0052	-.0009
	[-.02, .28]	[-.05, -.00]	[-.02, .03]	[-.02, .02]
<i>Factor shocks</i>				
SD(level shock)	.0589	[.0572, .0605]		
SD(slope shock)	.0814	[.0787, .0840]		
SD(curv. shock)	.1538	[.1483, .1599]		
Corr(level, slope)	-.7450	[-.7673, -.7208]		
Corr(level, curv.)	.2611	[.2189, .3043]		
Corr(slope, curv.)	-.4217	[-.4613, -.3809]		
<i>Measurement errors</i>				
$\sigma_w^2$	.0039	[.0039, .0040]		

Posterior means and 95% Bayesian credibility intervals (based on 2.5 and 97.5 percentiles of draws from the posterior distribution) in squared brackets for each model parameter. Estimates of risk sensitivity parameters are boldfaced if the credibility interval does not straddle zero.

their data set and model *only the level factor* seems to carry risk. This issue will be discussed further below (Section 2.4.1).

With regards to the shock covariance matrix  $Q = \Sigma\Sigma'$  standard deviations and correlations of the factor shocks are shown.<sup>22</sup> The shock covariance matrix is estimated rather precisely. There is strong correlation between the factor shocks. The measurement error variance is relatively small – it implies that the error is less than 12 basis points 95% of the time.

<sup>22</sup>An advantage of MCMC is that inference on non-linear functions of model parameters is just as precise as the inference on the original parameters. One simply calculates the parameters of interest (here standard deviations and correlations) from the fundamental parameters ( $\Sigma$  in this case) for each draw in the MCMC sample, which delivers a sample from the marginal posterior for these derived parameters.

### 2.3.3 Term premium estimates

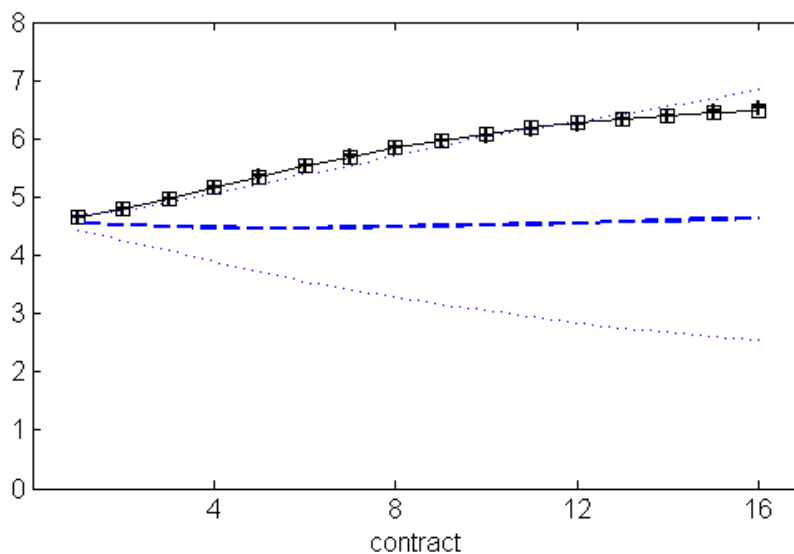
Our estimation results now allow us to decompose futures rates into risk-neutral rates and forward risk premia. First we consider average levels: Figure 2.1 shows for each of the 16 contracts the average level of the empirical and model-implied futures rates together with posterior means and 95%-CIs for the average level of the risk-neutral rates.<sup>23</sup> The posterior mean of the average risk-neutral rate curve is flat at about 4.5%, which indicates that the short rate was, on average in our sample, expected to remain about constant. The horizontal distance between risk-neutral rate curve and futures rate curve corresponds to the forward risk premium, and the point estimate for the average forward risk premium in the ED16 rate is about 200 basis points. This seems reasonable for a four-year forward risk premium in light of the existing evidence, see for example Kim and Orphanides (2007). However, the graph reveals that estimates of average risk-neutral rates, and hence of average forward risk premia, are extremely imprecise: The CI for the risk-neutral rate at the long end extends from about 2.5% to almost 7%. The large amount of uncertainty implies that the true average four-year forward risk premium could be anywhere between about -20 and +400 basis points.

How does the risk-neutral rate behave over time? Figure 2.2 plots for the ED16 contract time series of the fitted rates (these are indistinguishable from the actual rates) and the risk-neutral rates, for which point-wise posterior means and 95%-CIs are shown. The risk-neutral rate is much less variable than the actual rate and hovers around 3.5% to 5.5%. The resulting forward risk premium thus accounts for most of the variability of the actual futures rate. The actual rates have declined over the period in the sample, and based on the point estimate for the risk-neutral rate series it seems that this trend was entirely accounted for by a decline in the forward risk premium. But the CIs show a large amount of uncertainty underlying this point estimate. Looking at recent values for example, we cannot say whether the forward risk premium was +150 or -200 in the first

---

<sup>23</sup>For given values for the parameters and latent factors, I calculate the risk-neutral rates using equation (2.12) and average them across time. This is repeated for each draw in the MCMC sample, so that for each contract we have a sample from the posterior distribution of the average risk-neutral rate.



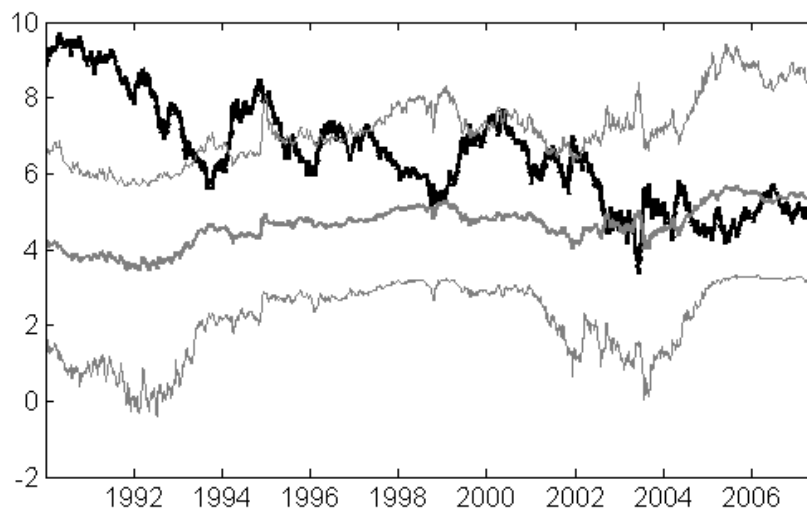


**Figure 2.1:** Average levels of actual, fitted and risk-neutral rates for unrestricted model

Average actual futures rates (crosses), average fitted futures rates (squares) and average risk-neutral rates (dashed line) with 95% credibility intervals (dotted lines). Units are percentage points.

half of 2007. In the words of Cochrane (2007, p. 278), “when a policymaker says something that sounds definite, such as ‘[...] risk premia have declined,’ he is really guessing”. The approach presented in Section 2.4 will be able to reduce this uncertainty.

Turning to high-frequency changes in risk premia, the focus of this paper, it is instructive to first consider some specific days with news events. We will consider four days, two with large positive payroll surprises (03/08/1996, +408,500 and 04/02/2004, +208,000) and two with monetary policy surprises (04/18/1994 and 03/22/2005). The effects of these news events on the term structure are best visualized by showing changes of actual futures rates across maturities, as well as changes in risk-neutral rates, which represent the revisions to short rate expectations in response to the news events. Figure 2.3 shows the actual and model-implied rate changes together with posterior means and 95%-CIs for the



**Figure 2.2:** Time series of fitted rates and risk-neutral rates for unrestricted model

Time series of fitted rates (black line) for ED16 contract and posterior mean of corresponding risk-neutral rate (thick grey line) together with 95% point-wise credibility intervals (thin grey line). Units are percentage points.

changes in risk-neutral rates across contracts. This is, to my knowledge, a new way to graphically analyze the effects of news events on the term structure of interest rates.

According to the point estimates changes in the short-maturity contracts are mostly due to changing short rate expectations, whereas for long maturities the *changes are attributed entirely to changing term premia*. Since revisions to short rate expectations are estimated to be close to zero at the long end of the term structure, changing forward risk premia alone account for the procyclical changes in long rates. Also evident from the figure is the *dramatic estimation uncertainty for changes in risk-neutral rates* – the CIs are very large. We cannot say with any confidence what happened to term premia in response to these news events when we appropriately account for the estimation uncertainty.

Figure 2.4 shows the empirical vol curve, i.e. the sample standard devia-

tions of daily futures rate changes, and the model-implied vol curve. Furthermore it shows posterior means and 95%-CIs for the risk-neutral vol curve. The well-known hump shape (Dai and Singleton, 2003) is clearly present in the vol curve of Eurodollar futures. The vol curve of only the unpredictable component of rate changes (not shown) is virtually indistinguishable from the actual vol curve, indicating that the predictable component of daily rate changes is very small.

What is the relative importance of variation in short rate expectations and forward risk premia for the volatility of futures rates? According to the posterior mean of the risk-neutral vol curve, while changes in risk-neutral rates alone drive the volatility at the short end, they account for only less than half of the volatility at the long end. Notably the estimation uncertainty for the risk-neutral vol curve is extraordinarily large.

The above results show two important problems with the term premium estimates of conventional, unrestricted DTSMs. The first one is obvious: The estimation uncertainty underlying estimates of levels and changes in risk-neutral rates is tremendous. This is due to the lack of precision in our estimates of the physical dynamics. I term this the “uncertainty problem”. My approach allows to quantify the uncertainty, which constitutes a challenge for most DTSMs but is hardly ever explicitly recognized.

The second issue is that the decompositions are implausible: The unrestricted model implies that forward risk premia account for the majority of rate changes at the long end of the term structure. The case studies showed that the entire procyclical response of long forward rates to the news events is attributed to forward risk premia. The risk neutral vol curve implies that for daily changes in long forward rates forward risk premia are a more important source of volatility than short rate expectations. However, we think that the term premium moves slowly, at business cycle frequencies, thus it should not account for a lot of daily movements of interest rates. Also we expect term premium movements to be countercyclical, hence its strong procyclical response to the news events comes as a surprise. The reason for the implausible decompositions is what I term the “bias problem”: Unrestricted DTSMs with stationary  $\mathbb{P}$ -dynamics imply a high esti-

mated speed of mean reversion for the short rate. Hence shocks die out quickly, far-ahead expectations of the short rate hardly move at all, and most variation in long rates is attributed to risk premia. But since the short rate is very persistent, the speed of mean-reversion is likely to be significantly over-estimated. The closer the largest autoregressive root of a time series is to one, the more pronounced is the downward bias in its estimate – see Kendall (1954) or, more recently, Jardet et al. (2009) and the references therein. Hence the close-to-zero long-run revisions cannot be taken at face value, since they are due to biased estimates of the physical dynamics.

The uncertainty problem and the bias problem have been recognized by other researchers, particularly by Duffee and Stanton (2004), Kim and Orphanides (2005) and Kim (2007). Before presenting a new statistical framework to overcome these problems, I now turn to the systematic response of short rate expectations and term premia to macroeconomic news.

### **2.3.4 The impact of macroeconomic news**

Studies of the response of the term structure of interest rates to macroeconomic announcements (Fleming and Remolona, 1999b; Gürkaynak et al., 2005b) have found strong procyclical responses, with a distinct hump shape and a significant sensitivity of far-ahead forward rates. The common approach is to regress changes in yields or forward rates on a measure of the surprise in the announcement, usually taken to be the difference between released and forecast values. Estimates of a DTSM can be used to assess how much of these responses are due to changing short rate expectations and changes in risk premia, respectively, by using model-implied changes in risk-neutral rates or risk premia as the dependent variables in these regressions. This approach is employed by Beechey (2007), who uses estimates of risk-neutral rates from Kim and Wright (2005) and finds that the forward risk premium in long forward rates responds strongly procyclical to macro news and seems to account for the majority of the total response of forward rates.

There are two important problems with this approach: First, it does not account for the uncertainty underlying estimates of risk-neutral rates and forward

risk premia. In this section I perform inference that appropriately incorporates this estimation uncertainty and show that Beechey’s point estimates cannot be taken at face value. Second, the results are driven by the fact that Kim & Wright’s decomposition leads to implausibly high variability of term premia (see Beechey’s figure 1), typical for DTSMs without or with only minimal restrictions on the prices of risk. Section 2.4 will present an approach that overcomes both of these problems and leads to different conclusions.

The term structure innovations under the physical measure on a given day can be calculated as  $v_t = X_t - \mu - \Phi X_{t-1}$ , based on a set of parameter estimates and values for the latent factors. In order to assess the impact of macro news on the term structure we simply project the factor innovations on measures of the macro surprises using the following system:

$$v_t = \alpha + \beta^{(1)} s_t^{(1)} + \dots + \beta^{(r)} s_t^{(r)} + \eta_t, \quad (2.15)$$

where  $s_t^{(1)}$  to  $s_t^{(r)}$  are scalars that contain the surprise component on day  $t$  for each of  $r$  different macroeconomic data releases,  $\alpha$  and  $\beta^{(1)}$  to  $\beta^{(r)}$  are  $k \times 1$  vectors of unknown parameters, and  $\eta_t$  is a vector of innovations.<sup>24</sup> Equation-by-equation least squares is efficient, despite the innovations  $\eta_t$  being correlated across equations, because the regressors are the same in each equation (Zellner, 1962). The resulting estimates  $\hat{\beta}^{(j)}$  show the response for each of the  $k$  innovations to a one unit surprise in release  $j$ . If a specific piece of news tends to always have a similar impact on the term structure, then this will be reflected by the value of the corresponding  $\hat{\beta}^{(j)}$  vector. The change in the risk-neutral rate for contract  $i$  caused by a unit surprise in a specific news release is  $\tilde{h}_i' \hat{\beta}^{(j)}$ , corresponding to the first term in equation (2.13).

The approach of Beechey amounts to calculating the innovations based on the DTSM’s point estimates for the parameters and smoothed estimates for the factors, and then performing the regression in equation (2.15) taken as given these innovations.<sup>25</sup> However, this ignores the fact that the physical innovations  $v_t$  are

<sup>24</sup>It is necessary to include all data release series in the regression in order to partial out the impact of releases that occur on the same day.

<sup>25</sup>Beechey regresses changes in implied risk-neutral rates and forward risk premia on the sur-

not known but instead estimated in a first step. We can appropriately account for this uncertainty using the posterior sample for parameters and factors, as detailed in the following algorithm:

1. Obtain parameters and latent factors for the current draw from the joint posterior.
2. Calculate the loadings  $h_i$  and  $\tilde{h}_i$  from the parameters, and the innovations from the factors and parameters using the fact that  $v_t = X_t - \mu - \Phi X_{t-1}$ .
3. For each of the  $k$  factor innovations, sample the regression coefficients, corresponding to the relevant equation of the system in (2.15), from the conjugate normal posterior.<sup>26</sup>
4. Calculate the predicted response of model-implied futures rates ( $h'_i \hat{\beta}^{(j)}$ ) and of risk-neutral rates ( $\tilde{h}'_i \hat{\beta}^{(j)}$ ) for each futures contract to each of the  $r$  different news releases.
5. Unless the end of the MCMC sample is reached return to step 1.

This provides a distribution of response coefficients for actual and risk-neutral rates to any of the  $r$  news releases. This distribution importantly takes into account the different sources of uncertainty: the first-step uncertainty underlying estimates of  $v_t$  as well as the second-step uncertainty from the regression analysis.

Figure 2.5 shows the results obtained for six different macroeconomic releases: Non-farm payroll employment, the unemployment rate, hourly earnings, Core CPI and Core PPI (Bureau of Labor Statistics, BLS) as well as retail sales (Department of Commerce). The surprise component in the data release is calculated as the difference between the actually released number and the value expected by the market, which is then standardized to have unit variance in order to make the different news releases comparable. To measure the market expectation I take

---

prise measure, which is equivalent to the approach outlined above since model-implied rate changes are driven by just  $k$  underlying term structure factors.

<sup>26</sup>I specify the prior for the regression parameters to be independently normal with mean zero and large variance. The prior for the error variance is taken to be inverse gamma.

the median market forecast, which is compiled by Money Market Services the Friday before the announcement. The figure shows the empirical responses of futures rates to macro news with 95% confidence intervals, the responses of model-implied rates, and the posterior means and 95%-CIs for the responses of risk-neutral rates. Note that the model satisfactorily captures the response of futures rates to the news – the model-implied responses are within the confidence intervals for the empirical responses in all cases.

The responses of risk-neutral rates to macro announcements, while significant at the short end of the term structure, are estimated to decrease to zero quickly with maturity for all six news releases. Thus at the long end of the term structure short rate expectations seem to not respond at all. This replicates the Beechey-result that the procyclical responses of long rates are mainly attributed to changes in forward risk premia. However, this neglects the two problems of uncertainty and bias in the estimates of the risk-neutral rates.

With regard to the uncertainty problem, the graphs reveal that because of the estimation uncertainty we cannot say with much confidence how strongly short rate expectations at long horizons respond to macro news: The CIs for the revisions caused by news are very large. The conclusion that “movements in term premia, not expected future short rates, account for most of the reaction” (Beechey, 2007, p. 2) is not warranted when we appropriately account for the uncertainty in the estimates of risk-neutral rates. The bias problem is reflected in the implausibly small responses of short rate expectations at the long end of the term structure.

## 2.4 Restrictions on prices of risk

An unrestricted DTSM has unsatisfactory implications for estimates of the term premium: The estimation uncertainty is dramatic, due to a lack of precision in estimates of the physical dynamics. Furthermore term premia have implausibly high volatility, since the short rate’s speed of mean reversion is over-estimated. Most DTSMs, such as the ones in Dai and Singleton (2002), Ang and Piazzesi (2003) and Kim and Wright (2005), suffer from these issues.

The remedy against both of these problems is to incorporate additional information to pin down the term premium, taking the form of either additional data or constraints on the model.<sup>27</sup> This paper suggests to impose constraints on the market prices of risk. In their absence the physical dynamics and the cross-sectional dynamics are estimated independently of each other, hence the no-arbitrage assumption does not restrict the estimates at all. However if we restrict prices of risk then the information in the cross section of interest rates, which pins down the risk-neutral dynamics very precisely, is brought to bear on our estimates of the physical dynamics.<sup>28</sup> The big question of course is: Which restrictions are reasonable?

Common practice in the term structure literature is to first estimate a DTSM without restrictions, and then in a second step to re-estimate the model by imposing zero restrictions on those risk sensitivity parameters which are insignificantly different from zero or have the largest relative standard errors.<sup>29</sup> There are several problems with this approach. Choosing restrictions based on individual standard errors ignores the off-diagonal elements in the covariance matrix of the estimates – a joint restriction is chosen without considering joint significance. A related problem (and probably the reason why it is uncommon to test joint restrictions on the parameters of risk) is that the MLE standard errors for the parameters of a DTSM are rather unreliable.<sup>30</sup> Moreover, the choice of a significance level required for inclusion of the parameter or of a cutoff for the relative standard error is necessarily arbitrary. Most importantly, alternative sets of restrictions lead to economically significant differences in results, as I will show below, and this approach offers no guidance on which set of restrictions is more credible.<sup>31</sup>

---

<sup>27</sup>Examples of studies that use additional data are Kim and Orphanides (2005), who include interest rate forecasts from surveys, and Campbell et al. (2009), who proxy for the price of risk using a dividend/price ratio.

<sup>28</sup>This fact has been noted for example by Kim and Orphanides (2005) and Cochrane and Piazzesi (2008).

<sup>29</sup>Among the numerous studies employing this approach are the influential papers of Duffee (2002), Dai and Singleton (2002), Ang and Piazzesi (2003) and Kim and Wright (2005).

<sup>30</sup>There are several reasons to doubt these standard errors: There are multiple local maxima, the asymptotic approximation might not be valid, and the numerical approximations to gradient and Hessian of the likelihood function are imprecise.

<sup>31</sup>Kim and Orphanides (2005) report that in the context of their DTSM some of the “different choices of parameters to be set to zero [...] exhibited economically significant quantitative



So far the literature has not developed an econometric approach to select restrictions on the prices of risk.<sup>32</sup> This paper provides a new framework for choosing plausible restrictions, which is based on Bayesian model selection.<sup>33</sup> One challenge is that there are many possible restrictions, which is overcome by first identifying plausible candidates using a new MCMC algorithm that involves latent indicator variables (Section 2.4.1). For the smaller set of candidate specifications we can then estimate the parameters and posterior model probabilities more precisely (Section 2.4.2). This shows the economic implications of each model (Section 2.4.3), and how much support each specification receives from the data. Since no single model clearly dominates all other candidates, I employ, in a third step, Bayesian model averaging (BMA) in order to perform inference that incorporates model uncertainty (Section 2.4.4).

The resulting estimates turn out to be both more precise and more plausible than those of the unrestricted model. In particular far-ahead short rate expectations are found to be significantly more variable, implying a slow-moving term premium. Furthermore the pro-cyclical response to news is found to be mainly due to changing short rate expectations, and there is no puzzling pro-cyclical term premium response. These results accord well with the conventional wisdom about bond risk premia.

The model-implied decompositions into expectations and risk premium components are independently verified by assessing forecast accuracy of the models (Section 2.4.5), as well as model-implied return-predictability (Section 2.4.6), with encouraging results.

### 2.4.1 Identifying candidate specifications

We would like to know which zero restrictions on  $\lambda_1$  are supported most strongly by the data – since this paper is concerned with the time variation in differences” (p.11).

---

<sup>32</sup>Cochrane and Piazzesi (2008) have taken the question of risk-price restriction more seriously, however they choose their restrictions based on very specific evidence on expected excess returns for a particular frequency and data set, thus their approach is not generally applicable.

<sup>33</sup>For review articles on the topic of Bayesian model selection see for example Kass and Raftery (1995) and Clyde and George (2004).

risk premia, we will leave the vector  $\lambda_0$  unrestricted. The problem of selecting a particular restriction is related to the variable selection problem in multivariate regression analysis: In both cases we can introduce a vector of indicator variables that summarizes which parameters are allowed to be nonzero. For the regression context Dellaportas et al. (2002) developed the method of “Gibbs variable selection” (GVS) which delivers a sample from the joint posterior distribution of the regression coefficients and indicators. I adapt this method to the context of DTSM estimation.

Let  $\gamma$  be a  $k^2 \times 1$  vector of indicator variables, each of which is equal to one if the corresponding element of  $\text{vec}(\lambda_1)$  is allowed to be nonzero. The goal is to obtain the joint distribution of  $(\gamma, \theta, X)$ . Since the conditional posterior of  $\gamma$  given  $\theta$  and  $X$  can be derived, block-wise M-H can be used to obtain draws from this distribution. To assess the plausibility of a joint restriction on  $\lambda_1$ , represented by a specific value of  $\gamma$ , say  $\bar{\gamma}$ , we can consider the posterior probability  $P(\gamma = \bar{\gamma})$ , which is easily estimated by counting the number of draws for which  $\gamma = \bar{\gamma}$ . The algorithm is developed in Appendix 2.6.3, and I will refer to it as GVS, although it shares only the idea of latent indicator variables with the original GVS algorithm.

The matrix  $\lambda_1$  has  $k^2$  elements thus there are  $2^{k^2}$  possible specifications, a large number even for the case of only a few factors. Thus a very large MCMC sample would be necessary to precisely estimate the posterior model probabilities of all models. However, those specifications with high posterior probability are likely to appear quickly in the GVS algorithm. Our goal is to identify the most promising specifications, and running the sampler for a limited number of iterations will achieve this goal.<sup>34</sup>

With the sample from the joint posterior for  $(\gamma, \theta, X)$  at hand I select those models with a Bayes factor in comparison to the most plausible model of at most 20. The Bayes factor is equal to the ratio of posterior model probabilities, and a value larger than 20 can be considered strong evidence against the model (Kass and Raftery, 1995). This leads to inclusion of six models, which together account for a total posterior model probability of 75.8%. To be clear: of the  $2^9 = 512$

---

<sup>34</sup>This point was made by George and McCulloch (1993).

possible specifications in the model space of the GVS sampler, the six most frequent specifications plus the unrestricted model are the only ones that the subsequent analysis will consider – from now on the model space will consist only of these seven models.

**Table 2.2:** Model specifications

Model (1)	Freq. (2)	Specification (3) (4) (5)			Eigenv. (6)	LR Rev. (7)	$P_{M_j}^{GVS}$ (8)	$P_{M_j}^{Lapl}$ (9)	$P_{M_j}^{Cand}$ (10)	$P_{M_j}^{RJ}$ (11)
$M_1$	0.0%	111	111	111	0.99	0	0.0%	0.0%	0.0%	0.0%
$M_2$	46.3%	010	000	000	1.00	0.95	61.1%	58.2%	56.1%	48.8%
$M_3$	9.4%	011	000	000	1.00	-0.34	12.4%	10.7%	13.9%	15.9%
$M_4$	8.4%	010	000	010	1.00	0.96	11.1%	9.4%	12.1%	18.2%
$M_5$	6.5%	000	000	000	1.00	1.00	8.6%	18.4%	12.9%	4.1%
$M_6$	2.9%	110	000	000	0.99	0	3.8%	2.4%	3.3%	6.3%
$M_7$	2.3%	011	000	010	1.00	-0.34	3.0%	0.9%	1.8%	6.6%

Alternative model specifications and estimated posterior model probabilities. Columns two: frequency of specification in GVS algorithm (cumulative: 75.8%). Columns three to five indicate which elements in the respective columns of  $\lambda_1$  are restricted (0) and unrestricted (1). Column six: largest eigenvalue of  $\Phi$ . Column seven: long-run revision of short-run expectations in response to a unit level shock. Columns eight to eleven: estimates of the posterior model probabilities based on rescaled GVS frequencies, Laplace approximation to marginal likelihood, Candidate’s estimate of the marginal likelihood, and Reversible-Jump MCMC.

Table 2.2 shows the candidate models. The unrestricted model is denoted by  $M_1$  and the restricted models  $M_2$  to  $M_7$ . The second column shows for the restricted models the frequency of each model in the GVS sampler – the unrestricted model is not visited by the sampler. Since these estimated model probabilities add up to 75.8% and not to 100%, I report the normalized numbers in column eight. Columns three to five indicate which elements in the respective columns of  $\lambda_1$  are restricted (0) and unrestricted (1), e.g. for model  $M_2$  only the element in the second row of the first column is unrestricted.

Importantly, the sample from the posterior distribution of  $\gamma$  allows us to perform inference on the determinants of time-variation in risk premia. Cochrane

and Piazzesi (2008) find, based on an analysis of excess returns, that only the price of level risk seems to vary. Since monetary policy mainly has slope effects on the term structure they conclude that it cannot be the risk associated with policy shocks that varies over time. We can assess whether this finding is supported by our sample. Time-variation in the price of level risk corresponds to the presence of non-zero elements in the first row of  $\lambda_1$ . The posterior probability of this hypothesis is estimated to be 17.3%, indicating that it is rather unlikely that level risk varies.<sup>35</sup> The price of slope risk varies if the second row of  $\lambda_1$  is non-zero, and the posterior probability of this hypothesis is estimated to be 89.3%. Thus, in contrast to Cochrane and Piazzesi (2008), I find that the price of slope risk and not the price of level risk seems to vary over time. Further analysis is necessary to reconcile this difference, which could be due to the sample choice, the frequency of the model, or the method of inference. My results suggest that the compensation for slope risk, and thus possibly for the risk associated with policy shocks, seems to play an important role for variation in excess returns and term premia.

## 2.4.2 Within-model simulation and posterior model probabilities

Having identified the candidate models I proceed by estimating each model individually. The purpose of this second step of my estimation approach is to estimate the parameters of each model more precisely. Their posterior distributions are needed for estimating posterior model probabilities by marginal likelihood methods, for a precise assessment of the different specifications' economic implications, and for constructing efficient proposals for the joint model-parameter sampling by means of RJMCMC (Section 2.4.4). The MCMC algorithm for estimating the restricted models closely corresponds to the one for the unrestricted model, with the only difference that the each element of  $\lambda_0$  and  $\lambda_1$  is sampled separately.

After performing within-model estimation, posterior model probabilities can be estimated based on marginal likelihood approximations. I use two different

---

<sup>35</sup>This is the relative frequency of draws with at least one element of the first row of  $\lambda_1$  being non-zero.

approximations, a Bartlett-adjusted Laplace estimator and a version of Candidate’s estimator, both of which are described in detail in DiCiccio et al. (1997).<sup>36</sup> The resulting estimates are given in column nine and ten of Table 2.2. Comparing these probabilities to the ones obtained from the GVS algorithm in column eight we see that they provide a similar ranking and weighting of the models: The unrestricted model is extremely unlikely, model  $M_2$  is strongly preferred,  $M_3$  to  $M_5$  receive about 10-20% probability each, and  $M_6$  and  $M_7$  are least likely. The correspondence between the results from GVS and from marginal likelihood approximations is reassuring – numerical differences are due to the above-mentioned imprecision of GVS and to the different approximations employed to estimate marginal likelihoods.

The unrestricted model’s posterior probability is estimated to be zero. Instead, strongly restricted specifications are supported by the data – none of the preferred models have more than three unrestricted elements in  $\lambda_1$ . The fact that *the data clearly supports tight restrictions on the prices of risk* is very plausible if we believe in the absence of arbitrage: Long rates should have some relation to expected future short rates, meaning that the physical and risk-neutral dynamics *should* be close to each other, but unrestricted prices of risk completely disconnect the two. Only when prices of risk are restricted does the no-arbitrage assumption have any bite. My results clearly speak in favor of such restrictions.

A crucial characteristic of the candidate models is whether they imply a stationary or an integrated short rate. The largest eigenvalue of  $\Phi$  for each model is shown in column six. If for a particular model this is unity, then the physical dynamics are non-stationary, the short rate contains a unit root, and far-ahead expectations of the short rate are affected by current factor shocks. On the other hand for stationary  $\mathbb{P}$ -dynamics the short rate is mean-reverting and far-ahead short rate expectations are constant (in the limit). Remember that  $\Phi^{\mathbb{Q}}$  has a unit

---

<sup>36</sup>The first approximation,  $\hat{C}_B^*$  in those authors’ notation, is a localized (i.e. volume-corrected) version of the Bartlett-adjusted Laplace estimator (see DiCiccio et al., 1997, Section 2.2). The second approximation,  $\hat{C}^C$ , a Candidate’s estimator, is based on a simple Kernel density estimate of the posterior distribution (see DiCiccio et al., 1997, Section 2.5). For the volume I use 5% in both cases and I estimate the mode by taking that parameter draw which maximizes the posterior.

eigenvalue, that  $\Phi = \Phi^Q + \Sigma\lambda_1$ , and that we impose non-explosive dynamics (no eigenvalue of  $\Phi$  can be larger than one). Hence whether the system is stationary or not depends entirely on the restrictions on  $\lambda_1$ : For unrestricted  $\lambda_1$  and for most restrictions the system will be stationary, however some strongly restricted specifications lead to non-stationary dynamics. Importantly, *most preferred models exhibit a unit root for the short rate*. In particular model  $M_2$ , which is strongly favored by the data, implies non-stationary physical dynamics. The support in the data for a stochastic trend in the short rate stands in stark contrast to the implication of the unrestricted model that the short rate relatively quickly reverts to its unconditional mean. While the true short rate process cannot literally have a unit root, since it is bounded from below and usually remains in a certain range, its largest root is certainly very close to one. My evidence suggests that in a DTSM of daily frequency an integrated specification approximates the true data generating process better than a stationary specification.

In those restricted models that contain a unit root shocks to the level factor lead to revisions of far-ahead short rate expectations. Column seven shows these long-run revisions,  $\lim_{h \rightarrow \infty} (E_{t+1} - E_t)r_{t+h}$ , in response to a unit level shock – details about this calculation are given in Appendix 2.6.4. The models have very different implications for this long-run revision: For  $M_2$ ,  $M_4$  and  $M_5$  the long-run revision is positive and close to or equal to unity, implying that far-ahead short rate expectations move about as much as far-ahead forward rates. In models  $M_3$  and  $M_7$  on the other hand the long-run revision is negative, i.e. the forward risk premium increases by more than one in response to a unit level shock. Evidently there is specification uncertainty about this important aspect of the model.

### 2.4.3 Economic implications of alternative models

The previous analysis has revealed that model  $M_2$  receives strong support from the data. What exactly does this model imply for the properties of short rate expectations and risk premia? Figures 2.6 and 2.7 graphically summarize the implications of the unrestricted model ( $M_1$ ) and the favorite model ( $M_2$ ).<sup>37</sup>

---

<sup>37</sup>Note that Figure 2.6 simply pulls together what was shown in Figures 2.3, 2.4 and 2.5.

The first panel of each figure shows the model-implied changes in actual and risk-neutral rates on 8 March 1996, where markets saw a strong positive payroll surprise of +408,500. The second panel shows the actual and risk-neutral term structure of volatility. The third panel shows the responses of actual and risk-neutral futures rates to a one-standard-deviation positive payroll surprise. Credibility intervals for the changes, volatilities and responses of risk-neutral rates indicate the estimation uncertainty.

Model  $M_2$  implies that *changes in short rate expectations play a dominant role for changes in actual futures rates* across the entire maturity spectrum. According to the point estimates for risk-neutral rate changes, the employment surprise in March 1996 caused futures rates to increase mainly because the expected future short rate path was revised upwards, and forward risk premia hardly changed on this day. The risk-neutral vol curve implied by  $M_2$  shows that most of the daily variability in futures rates results from changing short rate expectations. Furthermore model  $M_2$  attributes the pro-cyclical response of futures rates to payroll surprises mainly to revisions of the expected short rate path. The differences between models  $M_1$  and  $M_2$  are dramatic, with the implications of the latter model being more plausible in light of the conventional term premium wisdom.

The figure also reveals that the *decomposition obtained using the restricted model exhibits low estimation uncertainty*, indicated by the narrow credibility intervals. A policy-maker that believes in this model would not be guessing but instead could make confident statements about changes in policy expectations and risk premia. The strongly restricted prices of risk lead to high precision in the inference about short rate expectations.

If we were convinced that model  $M_2$  is the right specification, then we could stop here. However the other candidate models also receive some support from the data, hence we should consider their implications with regard to short rate expectations and risk premia. Figure 2.8 compares the models by showing in the first panel what happened on 8 March 1996, in the second panel actual and risk-neutral vol curves, and in the third panel the responses to payroll surprises. The models have identical implications for the changes in actual futures rates,

since the estimated risk-neutral dynamics are essentially the same for all models. However they differ significantly in their implications for risk-neutral rates.<sup>38</sup>

According to models  $M_2$ ,  $M_4$  and  $M_5$ , the volatility of futures rates across all maturities (second panel) as well as their pro-cyclical response to macro news (first and third panel) are mainly due to changes in short rate expectations. Since these models account for about 70-80% of the posterior model probability mass this constitutes evidence that far-ahead short rate expectations are quite variable, respond significantly to the news, and hence play an important role for determining rate changes.

However, models  $M_1$ ,  $M_3$ ,  $M_6$  and  $M_7$  imply that changes in forward risk premia account for most or all of the volatility and pro-cyclical responses of long rates – for the stationary models  $M_1$  and  $M_6$  this is due to the fact that shocks die out in the limit and do not affect far-ahead short rate expectations, whereas for the non-stationary models  $M_3$  and  $M_7$  the explanation is that a positive level shock is associated with a negative long-run revision. Since the candidate models have economically different implications, there remains some specification uncertainty.

#### 2.4.4 Accounting for model uncertainty using reversible-jump MCMC

When a model indicator is included as a parameter to be sampled using MCMC we speak of “joint model-parameter sampling.” Another approach than the product-space sampling introduced by Carlin and Chib (1995), of which GVS is a special case, is Reversible-jump MCMC, initially developed by Green (1995). This method is characterized by the ability to “jump” between models with parameter spaces of different dimensionality. Sampling simultaneously across model- and parameter-space using RJMCMC constitutes the third step of my estimation approach, the goal being to deal with specification uncertainty.

---

<sup>38</sup>The graph shows the posterior means for changes (first panel), volatilities (second panel) and responses (third panel) for the risk-neutral rates, as in previous figures. Calculating these objects of interest at the posterior means of the parameters leads to slightly different results (not shown) because of the non-linearity in the parameters – e.g. the posterior mean volatility is different from the volatility at the posterior mean. This aspect of estimation uncertainty is ignored in classical estimates of term premium characteristics.



A model indicator  $j \in \{1, \dots, J\}$  is included as an additional parameter, and the sampler, which is detailed in Appendix 2.6.5, approximates the joint posterior  $P(j, \theta_j, X|Y)$ . Because we have previously estimated all candidate models separately, we can use our previous results to choose very efficient proposal distributions.

Table 2.2 shows in the last column the estimated posterior model probabilities obtained using RJMCMC. The finding of model  $M_2$  being strongly favored by the data is confirmed, with its posterior probability estimated to be around 50%. The unrestricted model is never visited by the sampler; probabilities for the other models are estimated to be between 4% and 18%. These results support our previous conclusions.<sup>39</sup>

The sample obtained using the RJMCMC algorithm allows us to perform inference that accounts for specification uncertainty. This is crucial given the economically important differences between the models, and since we cannot be 100% sure that model  $M_2$  is the correct one, rather only about 50-60%, simply put. Bayesian model averaging (BMA) is the appropriate theoretical framework to incorporate specification uncertainty into our inference about risk premia: The posterior distribution of some object of interest, say the risk-neutral vol curve, conditional on only the data and not on a specific model, is obtained by averaging out the model indicator using posterior probabilities.<sup>40</sup> The RJMCMC sample conveniently delivers draws from the relevant posterior distribution of the parameters – the model indicator is “averaged out” if we simply ignore its value for each draw. I denote by *BMA* the model estimates obtained in this way.

Figure 2.9 shows the properties of the risk-neutral rates inferred using *BMA*. It corresponds to figures 2.6 and 2.7 except that now we do not condition on a specific model but instead average across models. The point estimates for changes and responses of risk-neutral rates and for the risk-neutral vol curve

---

<sup>39</sup>Note that the GVS and RJMCMC algorithms use the same priors and thus, despite the different model spaces, should deliver the same weightings of the candidate models. The differences indicate lack of convergence, the reason being, as mentioned above, that the large number of possible models in the GVS algorithm would require a huge number of iterations to achieve complete convergence.

<sup>40</sup>For an introduction to BMA see Hoeting et al. (1999).

obtained using *BMA* confirm the earlier conclusion: *Short rate expectations are the more important driving force for the daily volatility of the entire term structure and its pro-cyclical responses to macro news – forward risk premia move very little at the daily frequency.* This result stands in stark contrast to the implications of an unrestricted DTSM, where risk-neutral rates hardly move at all at the long end of the term structure and all daily variation is due to changing risk premia, as was shown in Section 2.3. The model in Kim and Wright (2005), being largely unrestricted, leads to the same implications, as shown by Beechey (2007), despite being augmented by survey forecasts. The decomposition of rate changes I obtain under sensible restrictions on the market prices of risk are more plausible: Since the term premium seems to move mainly at business cycle frequencies we would not expect it to account for much variability at the daily frequency. Furthermore the conventional wisdom, empirically and theoretically founded, has it that the term premium moves in a countercyclical way. Thus its contribution to the procyclical interest rate changes caused by news should be small. This is exactly what we find if we impose the restrictions on the prices of risk that are suggested by the data.

The previous conclusion was based on the point estimates obtained using *BMA*. Figure 2.9 also shows that averaging across models, instead of choosing a favored restricted model a priori, increases the uncertainty about short rate expectations and risk premia. This was of course to be expected. However, when comparing figures 2.6 and 2.9 it becomes evident that averaging across our restricted DTSM specifications leads to slightly *lower overall uncertainty* about changes in risk-neutral rates than for the unrestricted model, in particular when considering the risk-neutral vol curve. Restricting the prices of risk, even after accounting for model uncertainty, improves the precision of estimates of the short rate dynamics.

Another advantage of the approach developed in this section is that it solves the “discontinuity problem” documented by Cochrane and Piazzesi (2008) and Jardet et al. (2009): The implications of a DTSM for risk premia dramatically differ depending on whether the largest root of the short rate is equal to or less than unity. In reality this root is very close to but slightly less than one. A stationary model will underestimate this root as argued above, but a root of unity implied by

an integrated specification is not the literal truth either. The discontinuity problem is thus tantamount to the bias problem. Jardet et al. (2009) solve it using a “near-cointegrated VAR” based on the averaging estimators proposed by Hansen (2009). Effectively they average a stationary and a non-stationary specification. While this overcomes the discontinuity problem it does not solve the uncertainty problem – no restrictions are imposed on the prices of risk, so no-arbitrage is not brought to bear on estimation of the P-dynamics. The estimation uncertainty, which is not reported by the authors, is likely to be very large, as usual for unrestricted DTSMs. Using *BMA* on restricted specifications of a DTSM also amounts to averaging between stationary and non-stationary specifications, thus solving the discontinuity problem, but at the same time it solves the uncertainty problem. The result is that we obtain more precise and more plausible decompositions of interest rate levels and changes into expectations and risk premium components.

### 2.4.5 Forecast accuracy

Claiming that specific risk premium estimates are more plausible than others is equivalent to saying that the model more accurately captures the market’s short rate expectations. One way to evaluate this claim, based on the notion that market participants construct the best possible forecasts, is to assess the forecast accuracy of the model’s predictions for the short rate (e.g. Duffee, 2002; Dai et al., 2006).

As a simple reality check for the model’s forecasts, I construct in-sample forecast errors for the model-implied short rate  $r_t = X_t^{(1)} + X_t^{(2)}$ . The choice of in-sample forecasts is made for simplicity and data-availability – because of the parsimony of the restricted DTSMs we would expect these models to perform even better out-of-sample. I compare root mean squared forecast errors (RMSEs) across model specifications, including the forecasts based on *BMA*. To construct forecasts of the term structure factors, filtered values for  $X_t$  are used. As a point of reference, I include the RMSEs based on forecasts using a random walk for the short rate. Since we are mainly interested in long-horizon forecasts of several years, the horizons considered are 900, 1200 and 1500 days.

**Table 2.3:** Forecast accuracy

Horizon	RW	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$	$M_7$	BMA
900	3.00	2.25	2.19	2.09	2.19	2.28	2.08	2.09	2.11
1200	2.80	2.22	1.84	2.08	1.82	1.86	1.90	2.12	1.84
1500	2.66	2.14	1.61	2.05	1.60	1.60	1.80	2.13	1.65

Root mean squared errors for in-sample forecasts of the model-implied short rate, using alternative model specifications, compared to a random walk.

Table 2.3 shows the results. All models perform better than a random walk for the short rate. The restricted models produce superior forecasts of the short rate compared to the unrestricted model, particularly at longer horizons.<sup>41</sup> While it is not the case that our favorite model ( $M_2$ ) produces the best forecasts, this was to be expected, since the estimation and model selection procedures consider only one-step-ahead forecast errors. Almost without exception all restricted models beat the unrestricted model at all forecast horizons. Notably the models with stationary P-dynamics,  $M_1$ ,  $M_3$  and  $M_7$ , perform worse the longer the forecast horizon – the reason is that their forecasts are close to the unconditional mean of the short rate. The forecasts based on *BMA* perform very well, being a close second or third for each horizon. Since these are averages of the individual forecasts, their good performance reflects the generally good accuracy of forecast combinations (Timmermann, 2006).

A more detailed assessment of the forecasting performance of DTSMs with restricted prices of risk is warranted: The use of out-of-sample forecasts, rigorous inference about forecast accuracy, and inclusion of other forecasting methods will provide more detailed evidence as to whether this modeling approach can help improve interest rate forecasts. Because of the improved precision in estimates of the P-dynamics and the parsimony of the restricted models this seems to be a promising direction for future research.

The results above support the claim that the term premium estimates of

---

<sup>41</sup>The puzzling fact that forecast errors are smaller at longer horizon is due to the fact that the sample periods differ in each case.

restricted specifications are more accurate than those based on unrestricted prices of risk. Notably this analysis is about the *level* of short rate expectations and risk premia. The following analysis will assess the plausibility in terms of *changes* in risk premia, which have been the main focus of this paper.

## 2.4.6 Return predictability

In the presence of time-varying risk premia, changes in futures rates, like bond returns, are partly predictable. Using predictive regressions, Cochrane and Piazzesi (2005)(CP) find that one-year bond returns are explained by current forward rates with  $R^2$  of up to 44%, and similar results obtain for Eurodollar futures rates, as I will show below. On the other hand, the specification and parameter estimates of a DTSM have very concrete implications for the predictability of returns. This suggests that we can check the plausibility of model-based risk premium estimates by comparing the model's implications to the regression-based findings about return-predictability. In particular, the question is whether the restricted DTSMs imply similar predictability as we find in the data.

Which holding period should be considered? The predictable component of *daily* changes in futures rates is negligibly small, as mentioned above – daily changes are mainly driven by surprise changes in short rate expectations and risk premia. Inference about the predictable component of returns thus needs to be based on longer holding periods. For this reason we will consider one-year changes in Eurodollar futures rates, which correspond to absolute returns on positions in futures contract that are rolled over for four quarters and then liquidated. If I denote by  $N$  the length of this holding period<sup>42</sup> then the relevant return is  $ED_{t+N}^{(i)} - ED_t^{(i+4)}$ .

To assess predictability using return regressions, the one-year changes are projected onto current rates, using daily observations. Since the inclusion of all 16 futures contracts as explanatory variables leads to perfect multicollinearity, I restrict attention to only three contracts, namely ED4, ED9 and ED13 – these

---

<sup>42</sup>In the theoretical calculations this is taken to be equal to 260, the approximate number of business days in one year. In the data, where I take the holding period to be exactly one year,  $N$  varies between observations, a fact that the notation ignores for simplicity.

**Table 2.4:** Return predictability

Contract	data	$M_1$	$M_2$	$BMA$
ED4	.52 .46	[.11, .81]	.21 [.02, .52]	.27 [.02, .74]
ED8	.50 .44	[.12, .77]	.21 [.02, .51]	.25 [.02, .61]
ED12	.45 .43	[.14, .72]	.22 [.02, .52]	.25 [.03, .57]

$R^2$  in projection of annual returns on current rates/factors. For details please refer to text.

capture essentially all of the predictability.<sup>43</sup> Hence the regression specification is

$$ED_{t+N}^{(i)} - ED_t^{(i+4)} = \beta_0 + \beta_1 ED_t^{(4)} + \beta_2 ED_t^{(9)} + \beta_3 ED_t^{(13)} + u_t^{(i)}.$$

Table 2.4 shows in the first column the  $R^2$  for  $i = 4, 8, 12$ . A large share of the variance in returns, namely 45-52%, is predictable based on current futures rates. These numbers are in the ballpark of the results of CP.

In order to compare these regression-based results to the models' implications, I estimate the  $R^2$  based on simulated data for futures rates, assuming that the specific model is the true data-generating process. Specifically, for each set of parameters in the MCMC sample, I simulate time series for futures rates of length  $T = 4000$  (similar to the actual data), and run the same regressions as for actual futures rates. This leads to a sample from the posterior distribution for the  $R^2$  estimated using the typical regression approach.

The remaining columns in table 2.4 show the simulation-based  $R^2$  with 95%-CIs for models  $M_1$ ,  $M_2$  and  $BMA$ . The predictability in the simulated data is similar to what we found in the data.  $BMA$  implies simulation-based  $R^2$  of 25-27%, with CIs comfortably straddling the values found in the data. Hence risk premium estimates based on a DTSM with tight restrictions on the prices of risk estimates are plausible from the perspective of return regressions as well.

<sup>43</sup>These contracts are selected by choosing those three contracts with the highest predictive power for one-year changes, considering the average across contracts, i.e.  $\frac{1}{12} \sum_{i=1}^{12} (ED_{t+N}^{(i)} - ED_t^{(i+4)})$ . Including more contracts as explanatory variables in addition to ED4, ED9 and ED13 hardly changes the  $R^2$ .

## 2.5 Conclusion

This paper shows that conventional term structure models, which do not restrict market prices of risk, lead to unsatisfactory implications for short rate expectations and forward risk premia: The estimation uncertainty is too large for us to make any useful statements about bond risk premia. Furthermore term premia have implausibly high variability. These issues, which are particularly serious when we want to decompose changes at the daily frequency, are due to a disconnect between risk-neutral and physical dynamics. With unrestricted prices of risk the no-arbitrage assumption, which requires consistency between cross-sectional and dynamic properties of the term structure, does not restrict our estimates. I develop an approach that allows us to rigorously assess which restrictions on the market prices of risk are plausible. Estimation of restricted models brings in the information in the cross section to improve the precision of our estimates of the physical dynamics of the short rate. The inference about short rate expectations is thus much more precise than in a model without restrictions. The two main empirical results are: (1) The data supports tight restrictions on the prices of risk. (2) Under these restrictions short rate expectations, and not changing risk premia, account for the majority of daily volatility at the long end of the term structure and of the response to macroeconomic news. This contrasts with existing results and is more plausible in light of the conventional wisdom about the term premium.

A promising application of my framework is the context of macro-finance term structure models, which, as noted by Kim (2007), face the important challenge of putting more structure on the prices of risk: A key problem of these models is that the number of parameters is very large and the joint dynamics of term structure and macro variables are over-fitted. My approach can help overcome this problem since it greatly reduces the number of free parameters. In addition to leading to more parsimony and to more precise estimates of the factor dynamics, my framework allows researchers to perform rigorous inference on risk premia. The questions about which macroeconomic variables drive variation in risk premia and which macroeconomic shocks carry risk, described as “the Holy Grail of macro-finance” by Cochrane (2007, p. 281), can be answered by testing restrictions on

the prices of risk. The statistical framework I present in this paper allows to assess such restrictions, wherefore it is the right setting to tackle these questions.

## 2.6 Appendix

### 2.6.1 Change of measure

In order to show what kind of process the term structure factors follow under  $\mathbb{Q}$  we need to derive the conditional Laplace transform of  $X_{t+1}$  under  $\mathbb{Q}$ . We defined the one-period SDF, or pricing Kernel, as

$$M_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}\right),$$

which implies the change of measure based on the fact that for any one-period pricing Kernel we have<sup>44</sup>

$$M_{t+1} = \exp(-r_t) f^{\mathbb{Q}}(X_{t+1}|X_t) / f^{\mathbb{P}}(X_{t+1}|X_t).$$

Thus we obtain for the risk-neutral conditional Laplace transform

$$\begin{aligned} E^{\mathbb{Q}}(\exp(u'X_{t+1})|X_t) &= \int \exp(u'X_{t+1}) f^{\mathbb{Q}}(X_{t+1}|X_t) dX_{t+1} \\ &= \int \exp\left(u'X_{t+1} - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}\right) f^{\mathbb{P}}(X_{t+1}|X_t) dX_{t+1} \\ &= E\left[\exp\left(u'(\mu + \Phi X_t + \Sigma\varepsilon_{t+1}) - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}\right) | X_t\right] \\ &= \exp\left[u'(\mu - \Sigma\lambda_t + \Phi X_t) + \frac{1}{2}u'\Sigma\Sigma'u\right] \end{aligned}$$

which is recognized as the conditional moment-generating function of a multivariate normal distribution with mean  $\mu - \Sigma\lambda_t + \Phi X_t = (\mu - \Sigma\lambda_0) + (\Phi - \Sigma\lambda_1)X_t$  and variance  $\Sigma\Sigma'$ .

---

<sup>44</sup>The Radon-Nikodym derivative, which relates the densities under the physical and risk-neutral measure, is given by

$$\frac{f^{\mathbb{P}}(X_{t+1}|X_t)}{f^{\mathbb{Q}}(X_{t+1}|X_t)} = \left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)(X_{t+1}; \lambda_t) = \exp\left(\frac{1}{2}\lambda_t'\lambda_t + \lambda_t'\varepsilon_{t+1}\right).$$



Note that since  $X_t$  follows a Gaussian vector autoregression under  $\mathbb{Q}$  the model is in the  $DA_0^{\mathbb{Q}}(N)$  class of Dai et al. (2006).

The physical innovations  $\varepsilon_t$ , which are a vector martingale-difference sequence (m.d.s.) under  $\mathbb{P}$ , are related to the innovations under  $\mathbb{Q}$  by

$$\varepsilon_t^{\mathbb{Q}} = \varepsilon_t + \lambda_{t-1}.$$

Note that the risk-neutral innovations, while being m.d.s. under  $\mathbb{Q}$ , can have non-zero mean and be predictable under  $\mathbb{P}$ , depending on the prices-of-risk specification.

## 2.6.2 Basic MCMC algorithm and convergence diagnostics

### Likelihood functions

Denote by  $X$  the latent factors for all time periods, and by  $Y$  the full sample of observed futures rates. The likelihood of the factors is

$$\begin{aligned} P(X|\theta) &= P(X|\rho, \lambda_0, \lambda_1, \Sigma) \\ &= \prod_{t=2}^T (2\pi)^{-\frac{k}{2}} |Q|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} v_t' Q^{-1} v_t\right) \end{aligned}$$

where  $v_t = X_t - \mu - \Phi X_{t-1}$ . Note that  $\Sigma$  determines not only the factor covariance matrix  $Q = \Sigma\Sigma'$  but also affects the physical dynamics  $\mu$  and  $\Phi$  (see equation (2.6)). For the distribution of the observations  $Y$  conditional on the factors  $X$  we have the likelihood

$$\begin{aligned} P(Y|\theta, X) &= P(Y|\rho, \sigma_w^2, X) \\ &= \prod_{t=1}^T \prod_{i=1}^m (2\pi\sigma_w^2)^{-\frac{1}{2}} \exp\left(-\frac{(ED_t^{(i)} - a_i - h_i' X_t)^2}{2\sigma_w^2}\right). \end{aligned}$$

### Block-wise Metropolis-Hastings

The joint posterior distribution of the model parameters and the latent factors is proportional to the product of the likelihood function for the data, the

likelihood function for the factors, and the joint prior:

$$P(\theta, X|Y) \propto P(Y|\theta, X)P(X|\theta)P(\theta)$$

A block-wise Metropolis-Hastings (M-H) algorithm is used in order to obtain draws from this posterior distribution: At each iteration one draws from the full conditional posterior distribution for each block of parameters, conditional on the other parameter values. If this distribution is not known in closed-form, a M-H step is used in order to obtain the desired draw, otherwise we can directly draw from the conditional posterior (this is called a Gibbs step). The latent factors are drawn using the Filter-Forward-Sample-Backward (FFSB) algorithm developed by Carter and Kohn (1994). Iteratively drawing the blocks in this way leads to a sample which is approximately distributed according to the posterior  $P(\theta, X|Y)$ , which is the stationary distribution of the Markov chain (Chib and Greenberg, 1995).

Iterating on this block-wise algorithm, the first  $B$  observations are discarded (the burn-in sample) so that the effect of the starting values becomes negligible. Of the following iterations, only every  $s$ th draw is retained, so that the number of iterations necessary for a sample of size  $G$  is  $B + s \cdot G$ . For the basic MCMC algorithm used to estimate a single DTSM specification the configuration is  $B = 20000$ ,  $G = 5000$  and  $s = 40$ . These values result from a careful inspection of convergence plots under different configurations, given the restrictions of computational costs and memory constraints. Notably not more than several thousand draws can be saved since every draw contains not only the parameters but also  $T \cdot k$  values for the sampled paths of the latent factors.

Priors need to be specified for the parameters  $(\rho, \lambda_0, \lambda_1, \Sigma, \sigma_w^2)$ . While for the purpose of estimation they could be taken to be diffuse or improper this would lead to problems when we turn to model selection – posterior distributions tend to be much less sensitive to the choice of priors than Bayes factors (Kass and Raftery, 1995).<sup>45</sup> For example, improper priors lead to undefined Bayes factors. Furthermore, very diffuse but proper priors will lead to results that necessarily favor the restricted model (the Lindley-Bartlett paradox, see Bartlett, 1957). Given

---

<sup>45</sup>A Bayes factor is the ratio of the posterior model probabilities for two competing models/hypotheses.

the focus of this paper on selecting restrictions on  $\lambda_1$  (Section 2.4) this prior should not be too diffuse.

I specify  $\rho$  to be uniformly distributed over the unit interval. The prior for  $Q = \Sigma\Sigma'$  is Inverse Wishart (IW) and the prior for  $\sigma_w^2$  is Inverse Gamma (IG), both rather dispersed. The elements of  $\lambda_0$  and  $\lambda_1$  are normally distributed, independent, with mean zero and unit variance. The absolute magnitudes of the estimates for these parameters are small, thus despite the unit variance the prior is not very informative. Sensible alternative choices hardly affect the estimates I obtain. The joint prior  $P(\theta)$  also imposes the restriction that the eigenvalues of  $\Phi$  are at most one in absolute value, thus preventing explosive dynamics.

Instead of successively drawing every block in each iteration, one can randomize which block is sampled next, in which case we speak of Random Scan Metropolis-Hastings. I choose this method since one can fine tune how frequently each block is sampled: Those blocks are sampled more frequently which are more problematic in terms of mixing properties, and the blocks with parameters that mix well are sampled less frequently, which increases the efficiency of the algorithm. Specifically I sample only one block in each iteration, and the five blocks  $X$ ,  $\rho$ ,  $(\lambda_0, \lambda_1)$ ,  $\Sigma$  and  $\sigma_w^2$  are sampled with probability 10%, 20%, 50%, 10% and 10%, respectively. In the following I describe how each block is sampled.

*Drawing the latent factors ( $X$ )*

Given  $\theta$ , draws for the latent factors are obtained by means of the FFBSB algorithm developed by Carter and Kohn (1994): Kalman filtering delivers an initial time series of the factors, and then one iterates backward from the last observation and successively draws values for the latent factors conditional on the following observation. A detailed explanation of the algorithm can be found in Kim & Nelson (1999, chap. 8).

*Drawing the risk-neutral dynamics ( $\rho$ )*

The risk-neutral dynamics are given by  $\Phi^Q$ , since we impose  $\mu^Q = 0$ . The prices of risk are taken as given when drawing this block, so drawing  $\Phi^Q$  affects not only  $a$  and  $H$  but also the transition matrix of the physical dynamics,  $\Phi$ . The matrix  $\Phi^Q$  is completely determined by the root  $\rho$ , for which we have the following

conditional posterior

$$P(\rho|\theta_-, X, Y) \propto P(Y|\theta, X)P(X|\theta)P(\theta)$$

where  $\theta_-$  denotes all other parameters except  $\rho$ . Since we cannot sample directly from this distribution –  $\rho$  enters the density in a complicated way – we need to employ a M-H step. Since only proposal draws that are close to the value from the previous iteration have a chance of being accepted, a Random Walk (RW) step is the natural choice: In iteration  $g$  we draw the parameter according to  $\rho^{(g)} = \rho^{(g-1)} + \zeta_\rho t_4$ , a fat-tailed RW with  $t_4$  being a random variable with a t-distribution with four degrees of freedom, and  $\zeta_\rho$  being a scale factor used to tune the acceptance probability to be around 20-50%, which is the recommended range in the MCMC literature (see Gamerman and Lopes, 2006, p.196). Since the proposal density is symmetric for a RW step, the acceptance probability is given by

$$\alpha(\rho^{(g-1)}, \rho^{(g)}) = \min \left\{ \frac{P(Y|\rho^{(g)}, \theta_-, X)P(X|\rho^{(g)}, \theta_-)P(\rho^{(g)}, \theta_-)}{P(Y|\rho^{(g-1)}, \theta_-, X)P(X|\rho^{(g-1)}, \theta_-)P(\rho^{(g-1)}, \theta_-)}, 1 \right\}$$

For the case that the prior restrictions ( $0 < \rho < 1$  and non-explosive  $\Phi$ ) are satisfied – the acceptance probability is zero otherwise – this is simply equal to the ratio of the likelihoods of the new draw relative to the old draw, or one, whichever is smaller.

*Drawing the risk sensitivity parameters ( $\lambda_0$  and  $\lambda_1$ )*

In order to draw the risk sensitivity parameters, we recognize that for their conditional posterior distribution we have

$$\begin{aligned} P(\lambda_0, \lambda_1|\theta_-, X, Y) &\propto P(Y|\theta, X)P(X|\theta)P(\theta) \\ &\propto P(X|\theta)P(\theta), \end{aligned}$$

where  $\theta_-$  denotes all parameters except for  $\lambda_0$  and  $\lambda_1$ , since the likelihood of the data for given risk-neutral dynamics does not depend on the prices of risk. The parameters enter the likelihood for the latent factors in a highly non-linear fashion thus we cannot directly sample from the conditional posterior distribution. I tried both RW and Independence Metropolis proposals, and found the former to work

better in this context. If there are no restrictions imposed on  $\lambda_1$  then I draw  $\lambda_0$  and each column of  $\lambda_1$  separately. The innovation for the RW is then a  $k \times 1$  vector of independent  $t_4$ -distributed innovations (one could of course use a multivariate  $t$ -distribution). For the case that some elements of  $\lambda_1$  are restricted to zero, I draw each non-zero element of  $\lambda_0$  and  $\lambda_1$  separately, using a univariate RW with  $t_4$ -distributed innovations. The scale factors are adjusted in order to tune the acceptance probabilities. After obtaining the candidate draw, the restriction that the physical dynamics are non-explosive is checked, and the draw is rejected if the restriction is violated. Otherwise the acceptance probability for the draw is calculated as the minimum of one and the ratio of the likelihoods of the latent factors times the ratio of the priors for the new draw relative to the old draw.

*Drawing the shock covariance matrix ( $\Sigma\Sigma'$ )*

For the conditional posterior of  $\Sigma$  we have

$$\begin{aligned} P(\Sigma|\theta_-, X, Y) &\propto P(Y|\theta, X)P(X|\theta)P(\theta) \\ &\propto P(X|\theta)P(\theta), \end{aligned}$$

where  $\theta_-$  denotes all parameters except  $\Sigma$ , since by the absence of convexity effects the shock variances do not enter the arbitrage-free loadings and thus the likelihood of the data is independent of  $\Sigma$ . Since we need successive draws of  $\Sigma$  to be close to each other – otherwise the acceptance probabilities will be too small – independence Metropolis is not an option. I found element-wise RW M-H to not work particularly well. A better alternative in terms of efficiency and mixing properties is to draw the entire matrix  $\Sigma$  in one step. I choose a proposal density for  $\Sigma\Sigma'$  that is IW with mean equal to the value of the previous draw and scale adjusted to tune the acceptance probability, which is equal to

$$\alpha(\Sigma\Sigma'^{(g-1)}, \Sigma\Sigma'^{(g)}) = \min \left\{ \frac{P(X|\Sigma^{(g)}, \theta_-)P(\Sigma\Sigma'^{(g)}, \theta_-)q(\Sigma\Sigma'^{(g)}, \Sigma\Sigma'^{(g-1)})}{P(X|\Sigma^{(g-1)}, \theta_-)P(\Sigma\Sigma'^{(g-1)}, \theta_-)q(\Sigma\Sigma'^{(g-1)}, \Sigma\Sigma'^{(g)})}, 1 \right\}.$$

Here  $q(A, B)$  denotes the transition density, which in this case is the density of an IW distribution with mean A.

*Drawing the measurement error variance ( $\sigma_w^2$ )*

The variance of the measurement error can be drawn directly from its conditional posterior distribution, i.e. we have standard Gibbs-sampling for this step.

The reason is that conditional on the latent factors, the other parameters and the data, the measurement errors can be viewed as regression residuals, and the IG distribution is the natural conjugate prior. Since I impose the variance to be the same across the  $m$  measurement equations, the residuals from all measurement equations are pooled. The conditional posterior for  $\sigma_w^2$  is the natural conjugate IG distribution.

### Convergence diagnostics

After having obtained a sample using the described algorithm, convergence characteristics of the chain need to be checked, in order to verify that the draws are from a distribution that is close to the invariant distribution of the Markov chain. Differently put, the question is whether the draws that we obtain are from a chain that is mixing well.

A very simple and intuitive check of whether the chain is behaving well is to look at trace plots, i.e. plots of the successive draws for each parameter. In addition to this visual inspection, one can calculate several convergence diagnostics.<sup>46</sup> The autocorrelations of the draws for each parameter give a first indication of how well the chain is mixing. A commonly employed method to assess convergence, developed by Raftery and Lewis (1992), is to calculate the minimum burn-in iterations and the minimum number of runs required to estimate quantiles of the posterior distribution with a certain desired precision. Moreover one can diagnose situations where the chain has not converged, as suggested by Geweke (1992), by testing for equality of means over different sub-samples. Gelman and Rubin (1992) have suggested to run parallel chains from different starting values and to compare within-chain to between-chain variance, which is a simple and effective way to check for convergence. I have applied these and some other convergence checks in order to find out how many iterations are needed for approximate convergence and how the algorithm can be tuned in order to improve mixing. The general conclusion is that a lot of iterations are needed because  $\rho$  and the elements of

---

<sup>46</sup>For surveys on convergence diagnostics see Cowles and Carlin (1996) and Brooks and Roberts (1998).

$\lambda_0$  and  $\lambda_1$  traverse the parameter space only very slowly. This is a result of the small innovations in the RW proposals, which are necessary to obtain reasonable acceptance probabilities. Therefore I choose long burn-in samples ( $B = 20,000$ ) and a large number of iterations ( $G \cdot s = 200,000$ ). Under this configuration the graphs and diagnostic statistics indicate that the chain has converged.<sup>47</sup>

### 2.6.3 MCMC algorithm: Latent indicator variables

The algorithm developed here is based upon the ‘‘Gibbs Variable Selection’’ (GVS) method of Dellaportas et al. (2002), which is a special case of the product-space sampling of Carlin and Chib (1995). What is particular to GVS is that the models are nested. The idea of product-space sampling is rather simple: In each iteration we keep track of the parameters of all models, not only of those that are included in the current model. This implies that the dimensionality of the space we sample from remains the same across models, which allows standard block-wise M-H sampling, in contrast to RJMCMC where the dimensionality differs between models. Since the models are nested, the complete set of parameters is simply the entire  $\lambda_1$  matrix, in addition to the other model parameters,  $(\rho, \lambda_0, \Sigma, \sigma_w^2)$ , which are also assumed to be shared among models. Keeping track of all parameters then just means that  $\lambda_1$  always contains  $k^2$  non-zero elements, but when calculating the likelihoods conditional on a specific set of restrictions, only those elements of  $\lambda_1$  that are ‘‘switched on’’ according to  $\gamma$  are taken to be non-zero.

When we sample the elements of  $\lambda_1$ , conditional on  $\gamma$ , we need to distinguish whether a particular element is currently included in the model, and thus our draw informed by the data, or whether it is currently excluded. In the latter case the data is not informative and we sample from a ‘‘pseudo-prior’’ or ‘‘linking density’’, a concept introduced to the theory of Bayesian model selection by Carlin and Chib (1995). More precisely, the conditional posterior of an arbitrary element of

---

<sup>47</sup>There certainly remains room for improvement of the algorithm. In particular one could use methods for speeding up convergence, such as the hit-and-run algorithm, adaptive direction sampling, or simulated annealing (see Gamerman and Lopes, 2006, Section 7.4).

$\lambda_1$ , which I denote by  $\lambda_i$ , is given by

$$P(\lambda_i|\lambda_{-i}, \gamma_i = 1, \gamma_{-i}, \theta_-, X, Y) \propto P(X|\theta, \gamma)P(\lambda_i|\gamma_i = 1) \quad (2.16)$$

$$P(\lambda_i|\lambda_{-i}, \gamma_i = 0, \gamma_{-i}, \theta_-, X, Y) \propto P(\lambda_i|\gamma_i = 0) \quad (2.17)$$

where  $\theta = (\rho, \lambda_0, \lambda_1, \Sigma, \sigma_w^2)$  as before,  $\theta_-$  denotes all parameters in  $\theta$  other than  $\lambda_1$ , and  $\lambda_{-i}$  ( $\gamma_{-i}$ ) contains all elements of  $\lambda_1$  ( $\gamma$ ) other than  $\lambda_i$  ( $\gamma_i$ ). These conditional distributions parallel the ones in equations (9) and (10) of Dellaportas et al. (2002).

I assume prior conditional independence of the elements of  $\lambda_1$  given  $\gamma$ .

For the case that  $\lambda_i$  is currently included, we sample from (2.16). Note that the conditional posterior only depends on the latent factors  $X$  and not on the data  $Y$ , since all parameters that determine the likelihood of the data are in the conditioning set. The prior for each price of risk parameter,  $P(\lambda_i|\gamma_i = 1)$ , is taken to be standard normal. A difference between the DTSM context and GVS is that the conditional posterior in (2.16) is not known analytically. Hence we need to employ Metropolis-Hastings to obtain the draws. I use a fat-tailed RW proposal, with scaling chosen to tune the acceptance probability.

If  $\lambda_i$  is not currently included, i.e. if  $\gamma_i = 0$ , it is drawn from the pseudo-prior  $P(\lambda_i|\gamma_i = 0)$ . I choose this distribution to be normal with mean and variance given by the sample moments of the marginal posterior draws of  $\lambda_i$  for the full, unrestricted model. Carlin and Chib (1995) recommend to choose a distribution for the pseudo-prior close to the actual posterior, which for the elements of  $\lambda_1$  is likely to be similar between full model and restricted models.

The conditional posterior distribution of an element of the vector of indicators is of course Bernoulli and the success probability is easily calculated based on:

$$\frac{P(\gamma_i = 1|\gamma_{-i}, \theta, X, Y)}{P(\gamma_i = 0|\gamma_{-i}, \theta, X, Y)} = \frac{P(X|\gamma_i = 1, \gamma_{-i}, \theta) P(\lambda_i|\gamma_i = 1) P(\gamma_i = 1, \gamma_{-i})}{P(X|\gamma_i = 0, \gamma_{-i}, \theta) P(\lambda_i|\gamma_i = 0) P(\gamma_i = 0, \gamma_{-i})} \quad (2.18)$$

Since I use an uninformative prior, putting equal weight on  $\gamma_i = 1$  and  $\gamma_i = 0$ , the last term cancels out. Denoting the above ratio by  $q$ , the probability with which we draw  $\gamma_i = 1$  is given by  $q/(q + 1)$ .<sup>48</sup>

---

<sup>48</sup>A subtlety, which is ignored in the above notation, is that the joint prior  $P(\gamma, \theta)$  imposes



The MCMC algorithm used to produce a sample from the joint posterior for  $(\gamma, \theta, X)$  is again random-scan block-wise Metropolis-Hastings: In each iteration the block to be updated, either  $X$ ,  $\rho$ ,  $\Sigma$ ,  $\sigma_w^2$ ,  $\lambda_0$ , or  $(\lambda_1, \gamma)$ , is selected at random, then the parameters in the block are drawn from their full conditional posterior distribution. Only the last block needs further explanation, the others are updated exactly like in the full model. Conditional on  $\theta_-, X$  and the data,  $(\lambda_1, \gamma)$  is drawn as follows: First the elements of  $\lambda_1$  are updated conditional on the value of  $\gamma$  from the previous iteration. Second, the elements of  $\gamma_i$  are drawn conditional on  $\gamma_{-i}$ ,  $\theta$ ,  $X$  and the data. I implement two different versions of the algorithm: In the first version I update all elements of  $\lambda_1$  and  $\gamma$  in each step. In the second version I randomly choose to update only one pair  $(\lambda_i, \gamma_i)$ .

I run the algorithm with a burn-in sample of size  $B = 50,000$  and a sample size of  $G = 100,000$ , using every 5th draw from a longer chain. In order to get an idea of the convergence properties of this algorithm, I run several chains and make sure that the results are similar across chains. Since the first several models, which account for a large share of the posterior model probability mass, are similar across different runs of the chain and between the two algorithms, we can be confident that we have identified the specifications with high posterior model probabilities. The final results presented in the text are obtained from aggregating the samples for two runs of the first algorithm and two runs of the second algorithm, i.e. are based on a sample of size 400,000.

I assess whether the results on the plausibility of different restrictions on  $\lambda_1$  are reasonable given the sample from the posterior for  $\lambda_1$  for the unrestricted model. This turns out to be the case, based on individual credibility intervals, and in particular based on highest-posterior-density regions resulting from a normal approximation to this joint posterior. This is an important reality check for the algorithm described above.

As mentioned previously, an important issue in this context is the prior for  $\lambda_1$ . I performed additional sensitivity analysis, for example changing the prior

---

that the physical dynamics resulting from any choice of  $\gamma$  and  $\lambda_1$  can never be explosive. This is easily implemented in the algorithm: If including a previously excluded element would lead to explosive dynamics then I simply do not include it, i.e. set  $\gamma_i = 0$ , and vice versa.

variance of the elements of  $\lambda_1$  by orders of magnitude. My findings clearly show that the results of the model selection exercise remain robust for different choices of the priors.

#### 2.6.4 Long-run revisions to short rate expectations

The changes in far-ahead expectations of the term structure factors, using the eigendecomposition  $\Phi = VDV^{-1}$ , are

$$\begin{aligned} \lim_{h \rightarrow \infty} (E_{t+1} - E_t)X_{t+h} &= \lim_{h \rightarrow \infty} \Phi^h \varepsilon_{t+1} \\ &= V(\lim_{h \rightarrow \infty} D^h)V^{-1}\varepsilon_{t+1} \end{aligned}$$

which can only be non-zero if one of the eigenvalues is unity in absolute value. Since  $\Phi^Q$  has a unit eigenvalue associated with the level factor, and  $\Phi = \Phi^Q + \Sigma\lambda_1$ , if  $\Phi$  has a unit eigenvalue it will be associated with the level factor. In this case we have

$$\lim_{h \rightarrow \infty} (E_{t+1} - E_t)X_{t+h} = V \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^{-1}\varepsilon_{t+1}.$$

For the long-run revision of short rate expectations we thus obtain

$$\lim_{h \rightarrow \infty} (E_{t+1} - E_t)r_{t+h} = \delta'_1 V \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^{-1} \begin{pmatrix} \varepsilon_{t+1}^{(1)} \\ 0 \\ 0 \end{pmatrix},$$

where  $\varepsilon_{t+1}^{(1)}$  is the level shock.

#### 2.6.5 Reversible-jump Markov chain Monte Carlo

Including the model indicator as a parameter, the model is now parameterized as  $(j, \theta_j, X_j)$ . Since the latent variables carry over between models, we can simply write  $(j, \theta_j, X)$ . The algorithm I implement in order to obtain draws from the posterior distribution for  $(j, \theta_j, X)$  randomly chooses, in each iteration, between a “within-model” step, where the parameters of the current model are updated just like in the algorithm to estimate each model separately, and a “model-jump” step.

If a jump is attempted, first the candidate model indicator  $j'$  is chosen randomly, with equal probability for all models other than  $j$ . The question is now how to propose values for the parameters of model  $j'$ , denoted by  $\theta_{j'}$ . I decide for the models to share the parameters  $\rho$ ,  $\Sigma$ , and  $\sigma_w^2$ , denoted here by  $\theta_-$ , but not to share any elements of  $\lambda_0$  and  $\lambda_1$ . It might seem that the models naturally share  $\lambda_0$  and those elements of  $\lambda_1$  that are unrestricted in both models. However this version of the algorithm turned out to be the more efficient than to have the models share as many as possible parameters, mainly because the posterior distribution of some elements of  $(\lambda_0, \lambda_1)$  differs between models. To construct  $\theta_{j'}$  I take  $\theta_-$  from the current model together with the proposed values for  $\lambda_{j'}$ , by which I denote all non-zero elements of  $(\lambda_0, \lambda_1)$  in model  $j'$ . To propose values for  $\lambda_{j'}$  I take the normal approximation to the posterior distribution of  $\lambda_{j'}$ , which we have available from the within-model simulation.

The idea of reversible-jump MCMC is that reversibility is ensured by matching the dimensions between candidate parameter-vector and proposed parameter-vector. The acceptance probability for the proposed jump is given by the minimum of one and

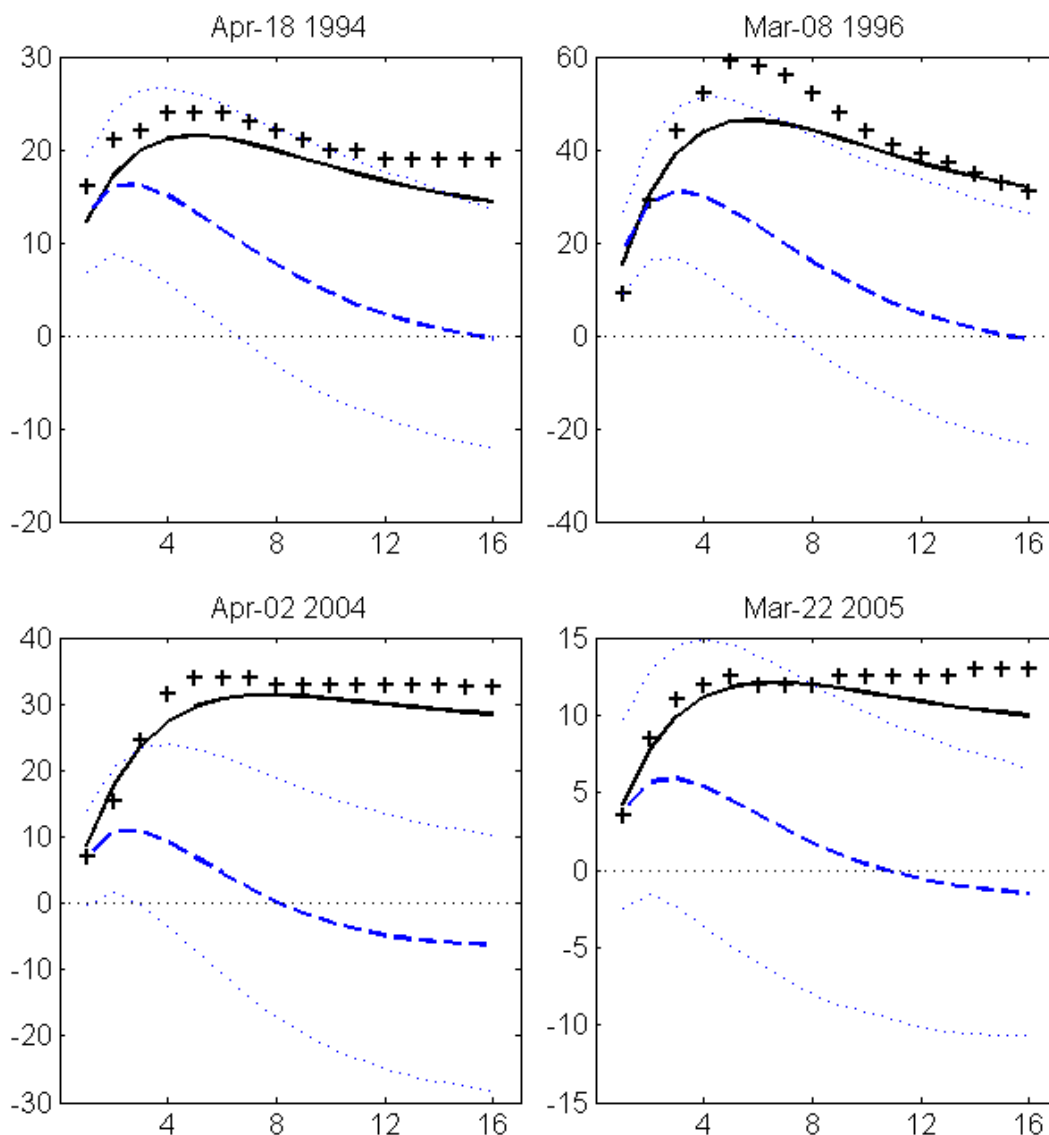
$$\frac{P(Y|j', \theta_{j'}, X)P(X|j', \theta_{j'})P(\theta_{j'}|j')P(j')}{P(Y|j, \theta_j, X)P(X|j, \theta_j)P(\theta_j|j)P(j)} \times \frac{q(u'|\theta_{j'}, j', j)q(j' \rightarrow j)}{q(u|\theta_j, j, j')q(j \rightarrow j')} \left| \frac{\partial g_{j,j'}(\theta_j, u)}{\partial(\theta_j, u)} \right|,$$

the product of model ratio (likelihood ratio times prior ratio) and proposal ratio. The parameter values for the candidate model are determined using  $(\theta_{j'}, u') = g_{j,j'}(\theta_j, u)$ , a bijection that ensures the dimension-matching. In our context  $(\theta_{j'}, u') = (\theta_-, \lambda_{j'}, u') = (\theta_-, u, \lambda_j) = g_{j,j'}(\theta_j, u)$  – the  $g$ -function is an identity function that simply matches the correct elements. Intuitively,  $u$  provides proposal values for all parameters in model  $j'$  that are not shared with model  $j$ , i.e.  $u = \lambda_{j'}$ , and  $u'$  takes on the values of the parameters in model  $j$  that are not used in model  $j'$ , i.e.  $u' = \lambda_j$ . Thus, recognizing the uniform prior over models, the equal jump probabilities ( $q(j \rightarrow j') = 1/6$  for all  $j \neq j'$ ), and the fact that the likelihood of  $Y$  given  $X$  only depends on  $\theta_-$  (which does not change between jumps) the above ratio simplifies to

$$\frac{P(X|j', \theta_{j'})P(\theta_{j'}|j')}{P(X|j, \theta_j)P(\theta_j|j)} \times \frac{q(u'|\theta_{j'}, j', j)}{q(u|\theta_j, j, j')}.$$

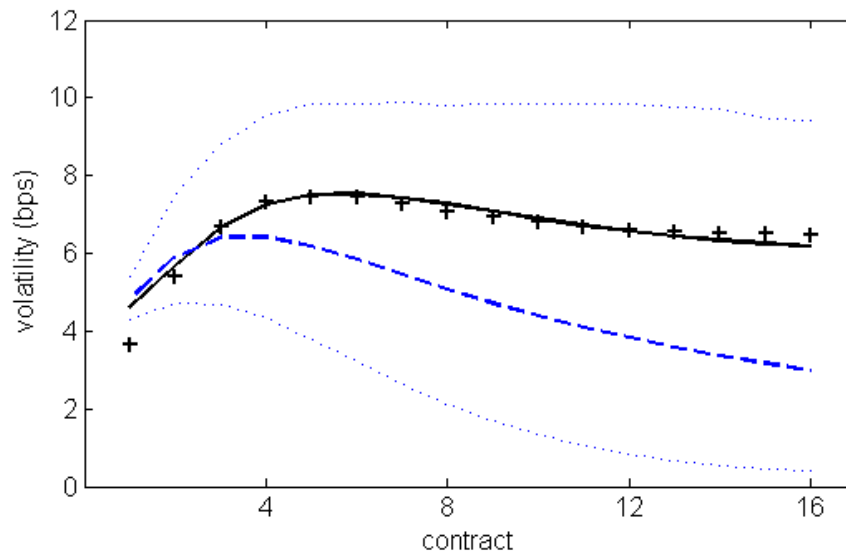
Note that since  $u' = \lambda_j$ , the distribution  $q(u'|\theta_{j'}, j', j) = q(\lambda_j|j)$  is the normal distribution with moments obtained from the sample from the posterior for model  $j$ , and correspondingly for  $q(u|\theta_j, j, j') = q(\lambda_{j'}|j')$ . Again, the minimum of the above expression and one is the probability with which we accept the proposed jump  $(j, \theta_j) \rightarrow (j', \theta_{j'})$ .

I run the sampler for  $B = 100,000$  burn-in iterations and then create a sample of length  $G = 5,000$  by using one out of every  $s = 200$  iterations. This is motivated by the fact that memory constraints make it impossible to save more draws of  $(j, \theta_j, X)$ , yet the sampler needs to be running for a considerable amount of iterations. Separate runs based on different starting values indicate that the chain has satisfactory convergence properties.



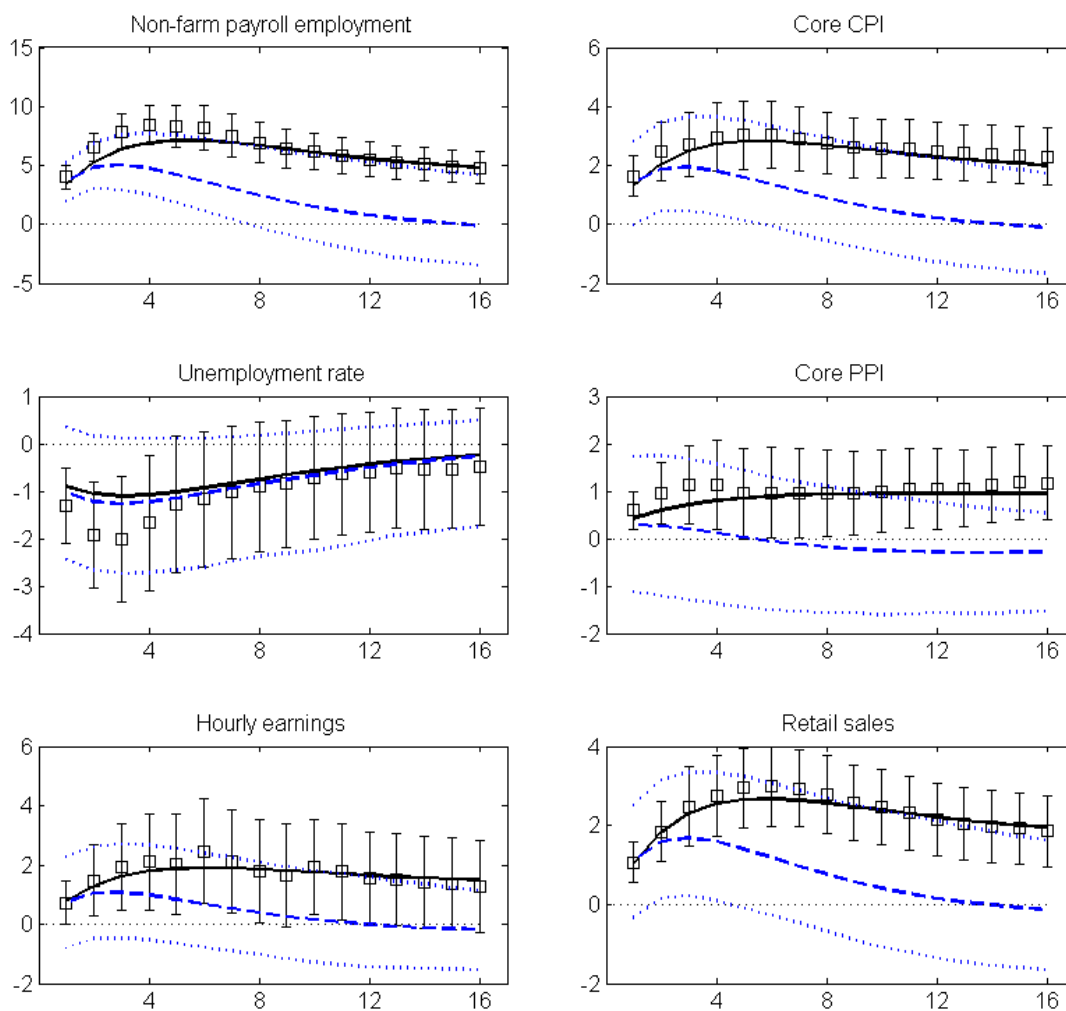
**Figure 2.3:** Effects of specific news events (unrestricted model)

Empirical (crosses) and model-implied (solid line) rate changes in response to certain news events, together with estimates of the revisions to short rate expectations: posterior means (dashed lines) and 95% credibility intervals (dotted lines) for changes in the risk-neutral rates. Units are basis points. Description of events: Apr-18 1994 – policy action, surprise tightening 25bps; Mar-08 1996 – payroll surprise, +408,500; Apr-02 2004 – payroll surprise, +208,000; Mar-22 2005 – policy action, surprisingly hawkish FOMC statement.



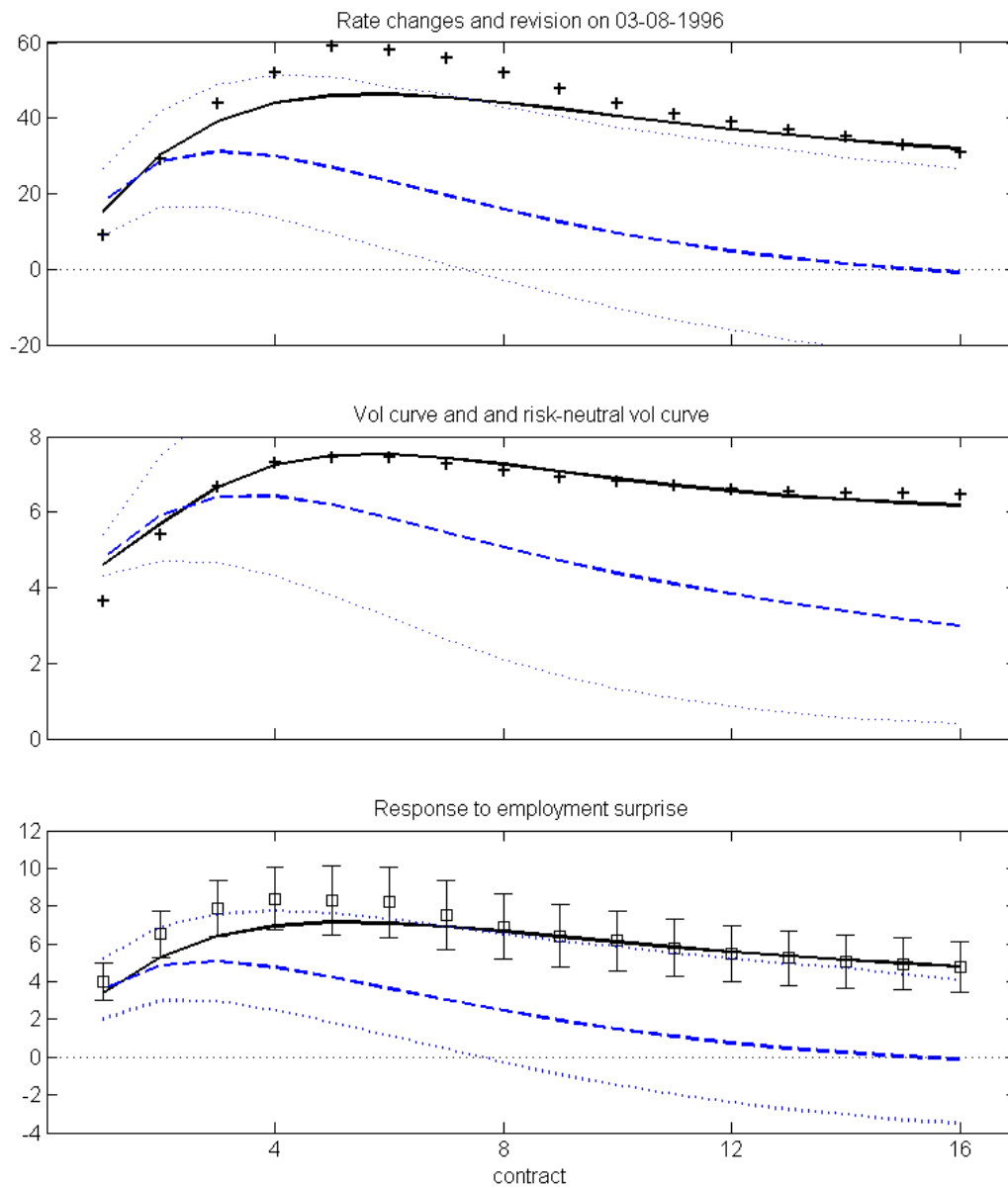
**Figure 2.4:** Term structure of volatility implied by unrestricted model

Estimates of actual and risk-neutral vol curve: Sample standard deviations (crosses) and model-implied standard deviations (solid line) of daily futures rate changes as well as posterior means (dashed line) and 95% credibility intervals (dotted lines) of standard deviations for risk-neutral rate changes. Units are basis points.



**Figure 2.5:** Responses to macro news implied by unrestricted model

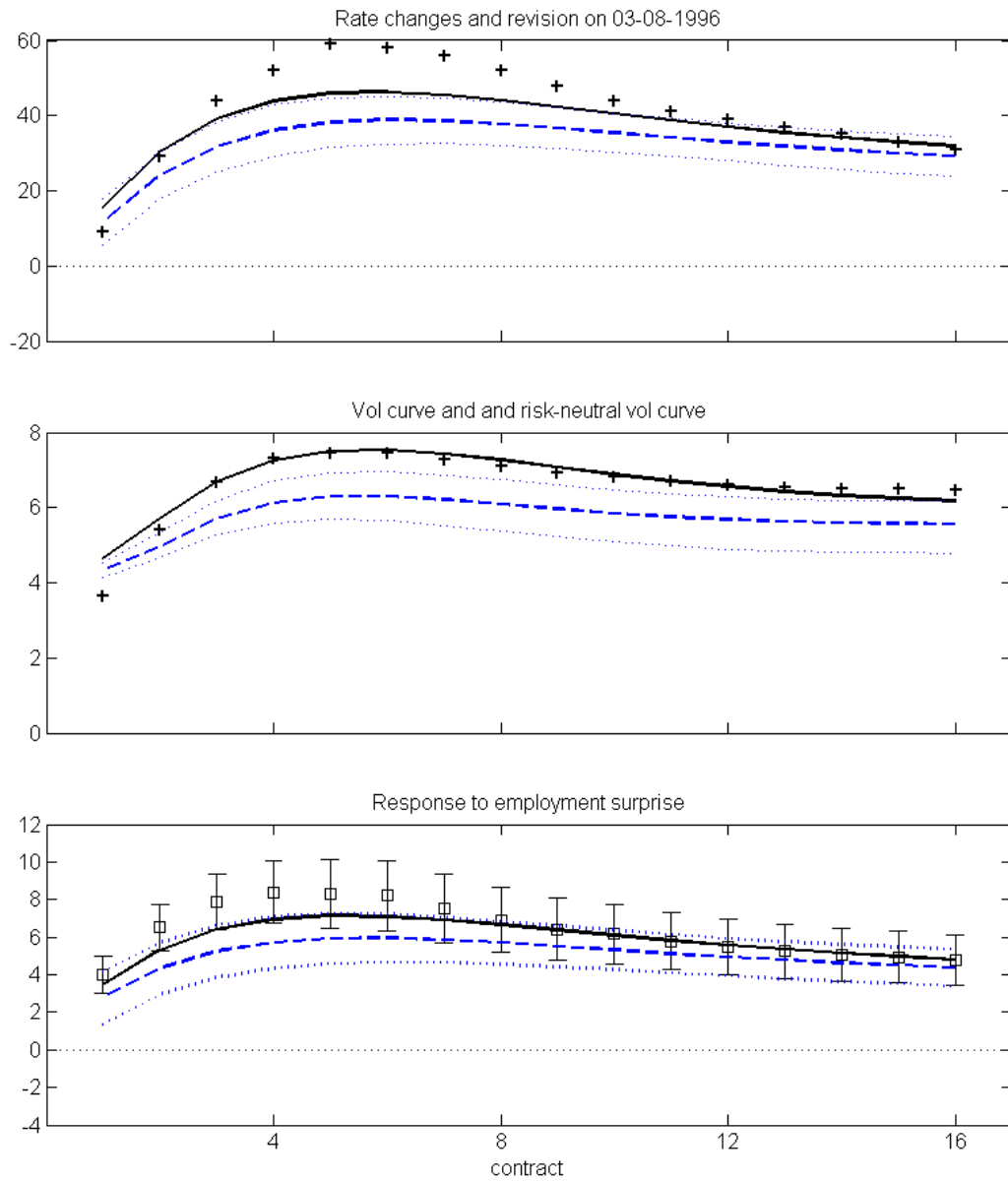
Responses to a one-standard-deviation surprise in six different macroeconomic data releases: Empirical responses of futures rates with 95% confidence intervals (error-bars), model-implied responses of futures rates (solid lines) as well as posterior means (dashed lines) and 95% credibility intervals (dotted lines) for estimated response of risk-neutral rates. Units are basis points.



**Figure 2.6:** Implications of unrestricted specification ( $M_1$ )

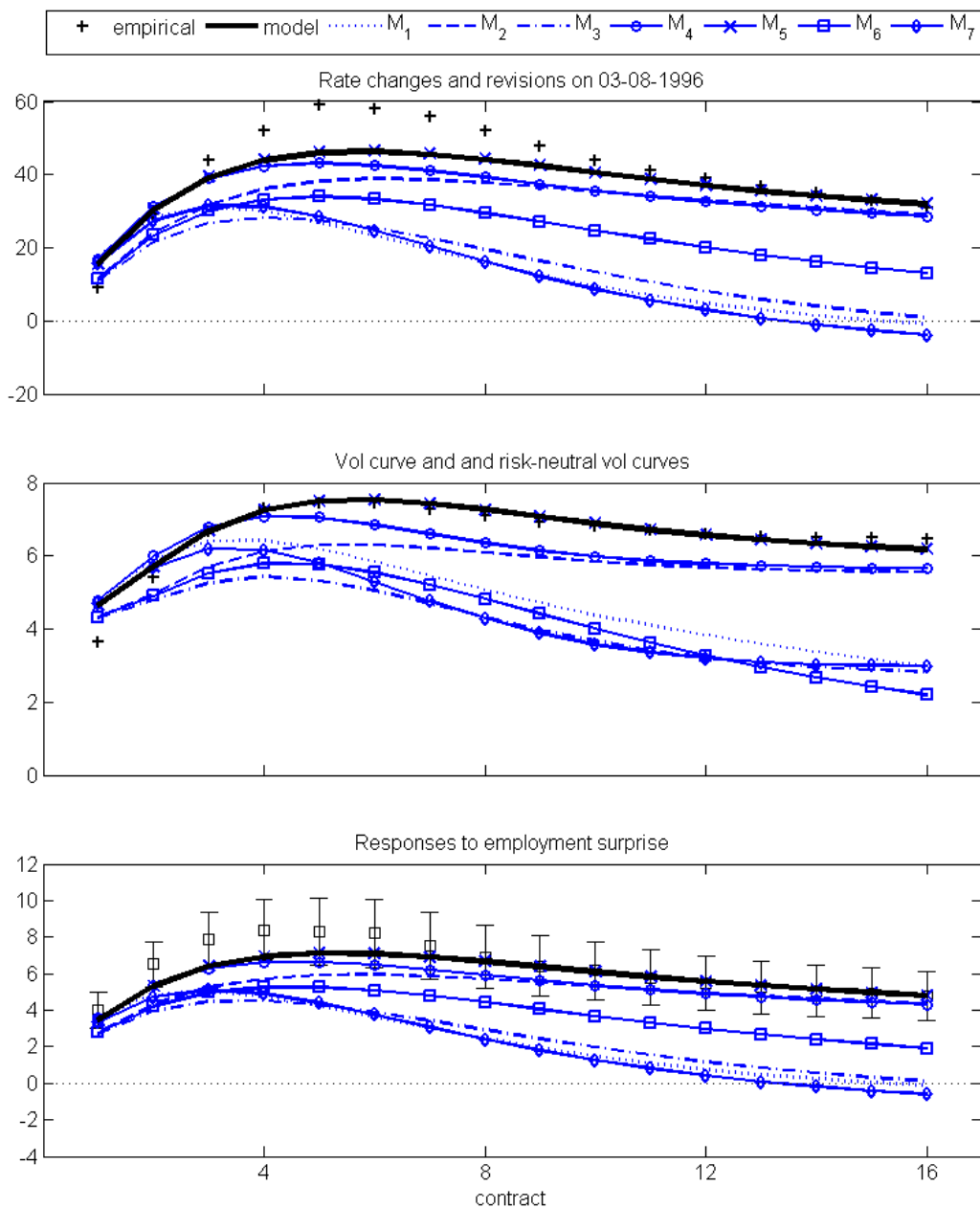
First panel: Empirical (crosses) and model-implied (solid line) rate changes on Mar-08 1996. Second panel: Sample standard deviations (crosses) and model-implied standard deviations (solid line). Third panel: Empirical responses of futures rates to a one-standard-deviation payroll surprise with 95% confidence intervals (error bars). Model-implied responses of futures rates to the news (solid lines). All panels show posterior means (dashed lines) and 95% credibility intervals (dotted lines) for the properties of risk-neutral rates. Units are basis points.





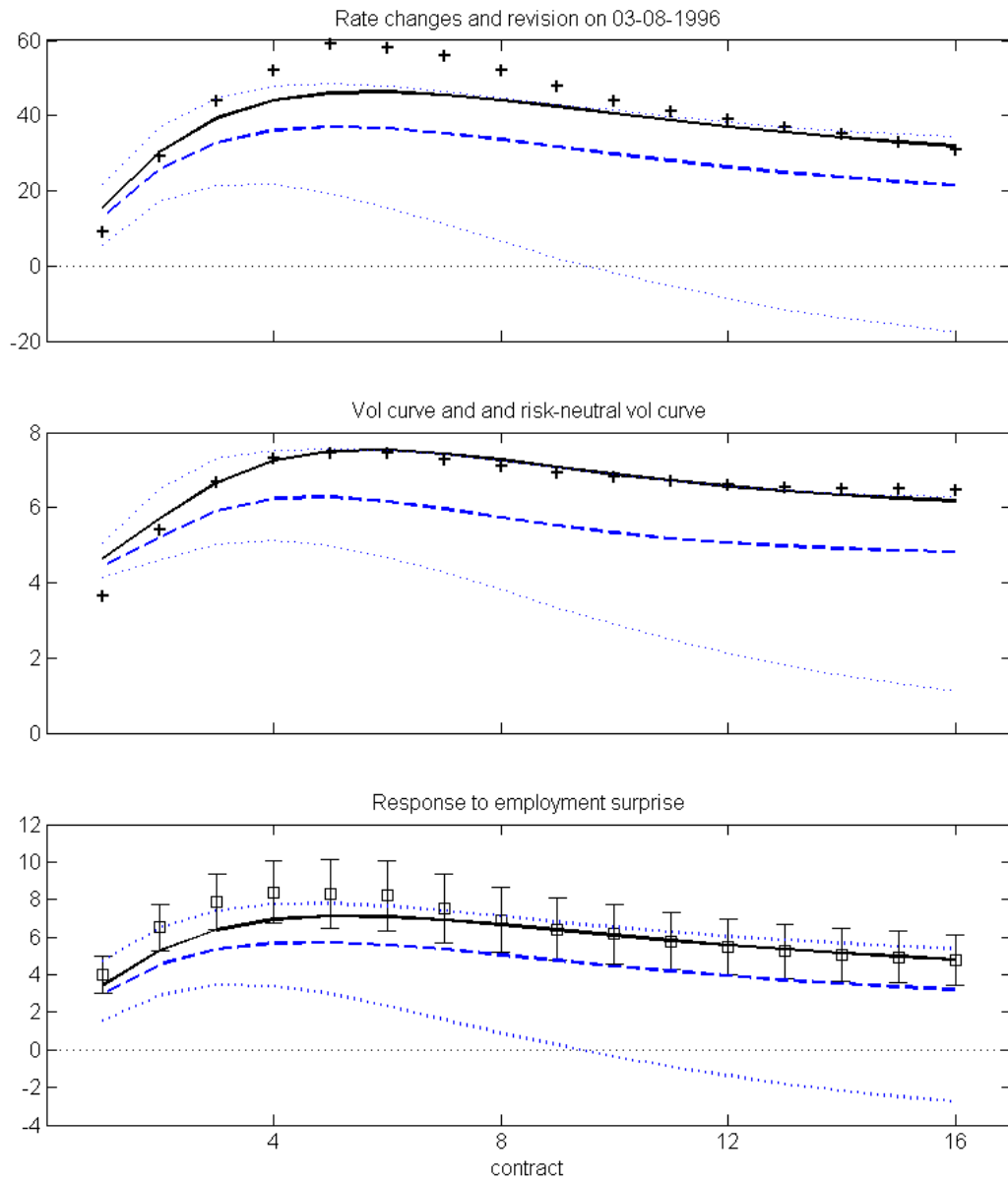
**Figure 2.7:** Implications of favored specification ( $M_2$ )

See description of Figure 2.6.



**Figure 2.8:** Comparison of alternative model specifications

First panel: Empirical (crosses) and model-implied (solid line) rate changes on Mar-08 1996 (payroll surprise +408,500), together with estimated changes in risk-neutral rates across models. Second panel: Sample standard deviations (crosses) and model-implied standard deviations (solid line) for futures rate changes as well as alternative risk-neutral vol curves. Third panel: Empirical responses of futures rates to a one-standard-deviation payroll surprise with 95% confidence intervals (error-bars), model-implied responses of futures rates to the news (solid line), and responses of risk-neutral rates implied by alternative model specifications. Units are basis points.



**Figure 2.9:** Implications of Bayesian model averaging (*BMA*)

See description of Figure 2.6.

## Chapter 3

# Inflation Expectations and the News

### Abstract

How do macroeconomic news change expectations of future monetary policy? Based on the Fisher equation, changes in expected future nominal rates can be decomposed into revisions of the expected path of the real short rate and changes in inflation expectations. Based on the evidence in the Taylor rule literature, which shows that changes in inflation account for a large portion of variation in policy rates, we would expect an important role for inflation expectations in explaining the effects of macro news on policy expectations. However the empirical evidence is not clear on this issue: some authors, using TIPS-based inflation expectations, find only minor effects of macro news on inflation expectations. This paper uses several alternative measures of revisions to inflation expectations, and provides evidence that surprises in macroeconomic data releases do in fact change inflation expectations quite significantly. This indicates that macro news change monetary policy expectations mainly by causing market participants to revise their expectations of future inflation.

### 3.1 Introduction

Nominal interest rates exhibit strong responses to surprises in macroeconomic announcements (Balduzzi et al., 2001b; Gürkaynak et al., 2005b). Chapter 2 has presented evidence that revisions of expected monetary policy, i.e. changes in expected nominal short term interest rates, are likely to be the main driving force for this phenomenon. Term premia are typically rather stable at these high frequencies and hardly respond to macro news. Based on the Fisher decomposition of nominal rates into real rates and expected inflation, the logical next question is to what extent expected real short rates and inflation expectations account for this response.

The Taylor rule literature has demonstrated that movements in policy rates can be accurately described by a simple policy rule in which the nominal short rate is a linear function of inflation and output.<sup>1</sup> In estimated Taylor rules the policy response to inflation is typically found to be stronger than responses to output (Taylor, 1999; Clarida et al., 2000). Based on this evidence one would expect that there exist a tight relationship between inflation expectations and short rate expectations, and furthermore that inflation expectations show a strong response to macroeconomic news and significantly contribute to explaining the response of nominal short rate expectations.

However some recent evidence by Beechey and Wright (2009), based on intra-day changes of inflation compensation, seems to indicate that inflation expectations hardly respond to macro news. Indeed the authors conclude that “the vast majority of the sensitivity is concentrated in real rates.” On the other hand there are studies (Gürkaynak et al., 2006a; Beechey et al., 2007) which, using daily data, find strong responses of inflation compensation. The evidence based on market-based measures of inflation expectations from TIPS data does not provide a satisfactory answer to the question of interest.

The purpose of this paper is to cast some more light on the question about the role of inflation expectations for changes in expected future monetary policy.

---

<sup>1</sup>After Taylor’s (1993) original paper there have been numerous studies on this topic, for a survey see Orphanides (2008).

The focus is to assess whether inflation expectations change significantly in response to real-side macro news, in particular to surprises in the non-farm payroll employment numbers. This data release is the one that leads to the strongest revisions of expected monetary policy and hence is a major source of interest rate volatility. Hence there is particular interest in understanding how the effects of this release decompose into inflation and real rate components.

To answer the question of interest I estimate revisions of inflation expectations using three alternative approaches: First, the common market-based measures are constructed using TIPS and nominal yield data. Particular attention is paid to differences between measures using intradaily and daily windows. Second, I use survey-based measures of inflation expectations. While these are only available at lower frequencies they can still be helpful to answer the question of interest. Third, I develop a novel approach to estimate revisions to inflation expectations based on macroeconomic news which is based on the concept of optimal linear forecasts, and essentially amounts to projecting future realizations of inflation on current macro surprises.

As it turns out the evidence speaks strongly in favor of an important role for inflation expectations in explaining variation in policy expectations. Not only do market-based measures of inflation expectations and short rate expectations show a close statistical relationship. In fact all three empirical approaches show that in response to real-side macro news inflation expectations are significantly revised. Furthermore the magnitude of these revisions is large enough to explain a sizeable share of the response of nominal short rate expectations.

The paper is structured as follows: Section 3.2 reviews the evidence in the Taylor rule literature about the role of inflation in explaining variation in policy rates, and provides some new estimates confirming large explanatory power of inflation in a simple univariate interest rate rule. A projection argument shows that revisions of short rate expectations should therefore be closely related to revisions of inflation expectations. Section 3.3 uses TIPS data and nominal yield data to construct proxies for these revisions and test this prediction. In Section 3.4 survey data is used to estimate revisions to inflation expectations, and their response

to macro news is estimated. Section 3.5 develops a novel empirical approach to estimate revisions to inflation expectations caused by macro news, based on future realizations of inflation after the release of macroeconomic data. Section 3.6 concludes.

## 3.2 Taylor Rules, Inflation Expectations and the News

In this section I first review some well-known evidence on the relationship between short-term interest rates and macroeconomic variables. Then I present some new evidence on how the short rate is related to measures of inflation in a univariate setting. Based on this evidence and on a projection of the policy rule on a past information set I state testable predictions about the relation between expected short rates and expected future inflation.

### 3.2.1 Taylor Rule evidence

A Taylor rule is a monetary policy rule that relates the policy rate to (past, current or expected future) output and inflation, and was introduced by Taylor (1993). In its most common form it is written as

$$i_t = r + \beta_\pi(\pi_t - \pi^*) + \beta_x x_t + u_t, \quad (3.1)$$

where  $i_t$  is the average policy rate in period  $t$ ,  $r$  is the equilibrium real rate,  $\pi_t$  is the inflation rate,  $\pi^*$  is the Fed's target for inflation, assumed to be time-invariant,  $x_t$  is the output gap, and  $u_t$  is a policy shock. The response coefficients  $\beta_\pi$  and  $\beta_x$  measure how the Fed reacts to inflation gap and output gap.

Taylor (1993) set  $\beta_\pi = 1.5$  and  $\beta_x = 0.5$  and showed that using this parametrization his rule provided an accurate description of actual monetary policy. Clarida et al. (2000) estimate a forward-looking version of the Taylor rule and find that for the Volcker-Greenspan era the inflation response coefficient is significantly larger than one. Ang et al. (2007) use term structure data together with macro data to estimate different Taylor rules, and confirm their large explanatory

power for interest rates. Hamilton et al. (2009) use the responses of federal funds futures and macroeconomic forecasts to macro news in order to identify Taylor rule parameters. They find large responses of the policy rate to inflation in such a market-perceived monetary policy rule. The broad picture that emerges from the Taylor rule evidence is that simple policy rules including measures of inflation and output have large explanatory power for short-term interest rates, and that real rates respond positively to variation in inflation during the Volcker-Greenspan era (the Taylor principle).

In order to derive predictions about the role of inflation expectations for changes in monetary policy expectations, I will estimate an interest rate rule where the nominal interest rate responds only to inflation. Since such a rule has only one regressor I refer to it as a “univariate Taylor rule”. Studies that employ such an interest rule specification include King (2000), Cochrane (2007b) and Del Negro et al. (2010). Such a univariate rule can be motivated by the empirically small response of nominal rates to the output variable (Taylor, 1999) as well as by the theoretical insight that in New-Keynesian macro models the welfare-optimal response to the output gap in simple interest rate rules is often close to zero (Rotemberg and Woodford, 1999). The empirical specification is

$$i_t = \alpha + \phi\pi_t + w_t, \quad (3.2)$$

with  $i_t$  denoting the average effective Federal funds rate during month  $t$  or quarter  $t$  – I will consider both frequencies –  $\alpha$  is an intercept, and  $w_t$  is a policy disturbance that can be serially correlated. The coefficient  $\phi$  measures how sensitive the Fed’s target is to inflation, and we typically speak of an “active” Taylor rule (or one that satisfies the Taylor principle) if  $\phi > 1$ . Estimating equation (3.2) will answer the following questions: (i) How did policy rates react to inflation in the past? (ii) How much variation in policy rates can be explained by variation in inflation?

Table 3.1 reports estimates of the response coefficient  $\phi$  in univariate Taylor rules based on different data sets. There are three dimensions across which the data sets differ: the price index used to calculate inflation, the frequency of the observations, and the interval over which price index changes are calculated. The price indices used are: the Consumer Price Index (CPI, reported by the Bureau of



Labor Statistics); the Core CPI, which excludes food and energy prices; the deflator for Personal Consumption Expenditure (PCE, reported by the Bureau of Economic Analysis); and the Core PCE. All price indices are seasonally adjusted. I consider monthly data as well as quarterly data. For monthly data I calculate inflation alternatively based on month-on-month (MoM), quarter-on-quarter (QoQ) and year-on-year (YoY) changes, and for the quarterly data set I consider QoQ and YoY changes. The table shows for each data set the estimate of  $\phi$ , its t-statistic based on Newey-West standard errors (12 lags for monthly data, 4 lags for quarterly data), and the  $R^2$ . The sample consists of the months from January 1987 (Greenspan became Chairman in 1987) to July 2007 (before the recent crisis started).

**Table 3.1:** Estimated univariate Taylor rules

Price index	Monthly ( $N = 247$ )			Quarterly ( $N = 82$ )	
	MoM	QoQ	YoY	QoQ	YoY
CPI	.77	1.00	1.25	.84	1.26
	(2.26)	(2.81)	(4.15)	(2.27)	(3.98)
	6.9%	16.0%	40.0%	13.0%	39.9%
Core CPI	2.72	2.54	1.42	2.49	1.45
	(4.40)	(5.57)	(4.57)	(5.37)	(4.67)
	27.9%	44.4%	47.2%	44.0%	49.1%
PCE	.96	1.20	1.23	.98	1.20
	(2.31)	(2.88)	(3.74)	(2.30)	(3.44)
	7.0%	15.9%	34.4%	12.1%	33.2%
Core PCE	1.84	2.13	1.36	1.81	1.34
	(2.83)	(4.35)	(4.13)	(3.48)	(3.98)
	14.4%	29.3%	39.2%	23.8%	38.3%

Estimates of  $\phi$  in a univariate Taylor rule, with t-statistics based on Newey-West standard errors and  $R^2$ , for different price indices, different frequencies, and different ways of calculating inflation – for details please refer to text.

The big picture is, not surprisingly in light of the evidence reviewed above, that inflation and the nominal short rate show a strong positive correlation, with  $R^2$  up to 50%, and that the response coefficient is generally above one.

The correlation is stronger if inflation is calculated using longer intervals. The reason is that monthly or even quarterly changes in a price index contain

much more noise, whereas comparisons with the previous year provide a cleaner picture of underlying inflation trends. The Fed of course does not want to react to noise but instead counter the true inflation pressures in the economy, thus the policy rate is more strongly correlated with those variables that better measure these inflation pressures.

For price indices which include food and energy correlations and response coefficients are smaller than for core price indices, because the latter reflects better the inflation tendencies that result from behavior of economic agents in the U.S. Interestingly, the response coefficients increase for the raw price indices going from month-on-month to year-on-year changes, whereas the response for core indices generally decreases the longer the change interval becomes. For YoY changes it hardly matters which price index we use: the response coefficients are all between 1.2 and 1.5.

To sum up the take-aways from the Taylor rule literature and from the above evidence: The coefficient on inflation in a Taylor rule is positive and generally larger than one, and inflation has high explanatory power for variations in the short rate.

### 3.2.2 Projecting a Taylor rule on past information sets

The evidence discussed above has some important implications for the relation between expected inflation and expected monetary policy, if we are willing to assume that market participants form expectations consistent with equation (3.2). To see this shift (3.2) forward  $h$  periods and take expectations with respect to the information set available to market participants at time  $t$ , with this conditional expectation denoted by  $E_t(\cdot)$ , to obtain

$$E_t i_{t+h} = \alpha + \phi E_t \pi_{t+h} + E_t w_{t+h}. \quad (3.3)$$

The idea of projecting Taylor rules on past information sets is due to Hamilton et al. (2009), who use the fact that revisions to short rate expectations are reflected in changes in Federal funds futures rates to identify Taylor rule parameters. We can now write

$$(E_t - E_{t-1})i_{t+h} = \phi(E_t - E_{t-1})\pi_{t+h} + (E_t - E_{t-1})w_{t+h}. \quad (3.4)$$

Hence if inflation forecasts are revised, policy expectations will also change, and the magnitude of the revision depends on  $\phi$ . Correspondingly, if we observe changes in monetary policy expectations, these can be due to either changing inflation expectations or changes in expectations of the future error term.

The large explanatory power of changes in inflation for variation in policy rates suggests that observed changes in monetary policy expectations are to a large extent due to revisions of inflation expectations. Furthermore the fact that estimates for  $\phi$  are usually larger than one implies that the sign of the revisions of inflation expectations would typically be the same as the sign of the change in expected policy rates, with the magnitude of the former close to but below the magnitude of the later.<sup>2</sup>

How can we test these predictions? The first issue is how to measure revisions of short rate expectations and inflation expectations. Based on nominal interest rates (from Treasury bonds) and real interest rates (from TIPS bonds), using the assumption that risk premia move slowly, we can construct good proxies for the two objects of interest. Alternatives are to use survey data or statistical forecasts to construct revisions – more on this below.

The second issue is which empirical methodology to use in order to test these predictions. One way of course is to simply regress a measure of the revisions of short rate expectations on a measure of the revisions of inflation expectations, and this will be done in Section 3.3.1 using market-based measures of revisions (i.e. using TIPS and nominal yields). In addition to an unconditional analysis the question will also be addressed whether the predictions hold in different subsamples, namely for days with particular news such as payrolls vs. CPI.

Another approach is to consider the response of inflation expectations to macro news, i.e. to surprises in macroeconomic data releases. We know that some releases have strong systematic effects on short rate expectations, such as the non-farm payroll numbers (Chapter 2). In light of the above discussion we would thus

---

<sup>2</sup>Note that this prediction is for “typical” revisions, since we are talking about population moments: there might be some news (changes in information sets) that have effects on expectations that are not in line with this prediction. But given the close relationship between inflation and short rates these atypical news should be rare.

expect this news to also significantly affect inflation expectations.<sup>3</sup> Furthermore we would expect the response to be of the same sign as the response of short rate expectations. I will test this prediction using three alternative measures for inflation expectations: inflation compensation from TIPS data (Section 3.3.2), survey-based inflation expectations (Section 3.4) and optimal linear inflation forecasts based on actual future inflation (Section 3.5).

### 3.3 TIPS Yields and Inflation Compensation

Treasury Inflation-Protected Securities (TIPS) are bonds that deliver payoffs indexed to the CPI, therefore yields correspond to real rates.<sup>4</sup> The difference between nominal rates and real rates for the same maturity is called “break-even inflation” or “inflation compensation” (IC), which equals the sum of expected inflation and an inflation risk premium. While this paper does not address the issue of estimating the inflation risk premium, most macroeconomists have the prior that risk premia move slowly, at business cycle frequencies, and Chapter 2 shows evidence that this is a reasonable assumption for the term premium. Thus changes in nominal rates approximate well the revisions of nominal short rate expectations. I will proceed under the assumption that the inflation risk premium also moves slowly, such that changes in inflation compensation can be taken to be good measures of revisions of inflation expectations.

#### 3.3.1 Correlation between nominal rates and inflation compensation

How are revisions of monetary policy expectations related to revisions of inflation expectations? One way to address this question is simple regression analysis. Consider the following regression, where  $t$  indexes business days:

$$f_t^n - f_{t-1}^n = \delta(fic_t^n - fic_{t-1}^n) + \varepsilon_t. \quad (3.5)$$

---

<sup>3</sup>Note that payrolls are the most important source of interest rate volatility (Chapter 1) so its effects on inflation expectations should be close to the unconditional predictions.

<sup>4</sup>For a detailed discussion of the TIPS market see Gürkaynak et al. (2010).

$f_t^n$  is the instantaneous nominal forward rate with maturity  $n$  years on day  $t$ , and the data are taken from Gürkaynak et al. (2006b). Hence  $f_t^n - f_{t-1}^n$  is a proxy for the revision on day  $t$  of the expectation about the nominal short rate  $n$  years ahead. The corresponding forward inflation compensation,  $fic_t^n$ , is calculated by subtracting the instantaneous real forward rate from the nominal forward rate, and the real rate data are provided by Gürkaynak et al. (2010).  $fic_t^n - fic_{t-1}^n$  proxies the revision on day  $t$  of the instantaneous  $n$ -year-ahead inflation rate. My sample starts in January 2003 (at which point the TIPS market had left its infancy) and ends in July 2007 (before the end of the recent turmoil). Estimating this regression for the ten-year maturity I obtain  $\hat{\delta} = .94$  with a standard error of 0.03 and  $R^2 = 42.8\%$ . This means that at a daily frequency, nominal *forward rates essentially move one-for-one with inflation compensation*, and that this explains about about 43% of the variation in nominal rates. The correlation is estimated to be  $\sqrt{.428} \approx 0.65$  and an approximate 95%-confidence interval<sup>5</sup> is given by [0.62, 0.69]. Clearly *changes in nominal forward rates are closely associated with changes in inflation compensation*. This is in line with our expectations: the prediction of a close association between revisions of short rate expectations and inflation expectations is confirmed by the large  $R^2$ . And while we predicted the coefficient to be above one, an errors-in-variables bias might explain the fact that it is estimated to be slightly smaller than one.

Is the relation between nominal rates and IC similar on days with employment reports compared to days with other news? Table 3.2 shows the results of estimating equation 3.5 in different subsamples: days with an employment report, days with a CPI report, and all other days. I report response coefficients and OLS standard errors, as well as correlations with 95%-confidence intervals. For all subsamples the response coefficients are large and close to unity, and the correlations between nominal rates and IC are sizeable. On days when the CPI is released, nominal rates response more than one for one to inflation compensation, corresponding to the Taylor principle.

Another question is whether the relation between nominal rates and infla-

---

<sup>5</sup>Take  $r$  to be the empirical correlation. I calculate a confidence interval for  $r$  based on the fact that the transformed random variable  $\zeta = .5 \log \frac{1+r}{1-r}$  is approximately normally distributed.

**Table 3.2:** Correlations across subsamples

Mat.	Response		Correlation	
<i>Employment report</i>				
5	0.64	(0.18)	0.43	[0.19, 0.63]
7	0.68	(0.15)	0.51	[0.28, 0.68]
10	0.90	(0.11)	0.73	[0.58, 0.84]
15	1.08	(0.13)	0.71	[0.54, 0.82]
<i>CPI report</i>				
5	1.41	(0.27)	0.58	[0.37, 0.73]
7	1.30	(0.16)	0.75	[0.60, 0.84]
10	1.21	(0.13)	0.79	[0.67, 0.87]
15	1.27	(0.14)	0.77	[0.63, 0.86]
<i>Other days</i>				
5	0.70	(0.04)	0.45	[0.40, 0.50]
7	0.95	(0.04)	0.57	[0.52, 0.61]
10	0.92	(0.03)	0.64	[0.60, 0.68]
15	0.85	(0.04)	0.59	[0.55, 0.63]

Relation between rate changes in nominal forward rates and changes in forward inflation compensation across different regimes: response of nominal rates to IC in univariate regressions (OLS standard errors in parentheses) and correlations (with 95%-confidence intervals in squared brackets) for days with employment reports, days with CPI reports, and other days.

tion compensation documented above is special to the daily frequency. It could certainly be the case that while at high frequencies there is a strong relation, for monthly or quarterly changes in nominal rates the driving force is variation in real rates, with the inflation compensation being more stable or unrelated to nominal rate changes. Table 3.3 shows estimates for the response coefficients and correlations for different frequencies: daily, weekly, monthly and quarterly changes. Importantly, the close association between nominal rates and the IC holds up across all frequencies. In fact if we go from daily to lower frequencies, the response coefficients are even higher and the correlations even stronger.<sup>6</sup>

<sup>6</sup>Note that for lower frequencies changes in risk premia come into play, thus rate changes are less accurate measures of the underlying changes in expectations. The higher response coefficients and  $R^2$  thus partly reflect the comovement of term premia and inflation risk premia.

**Table 3.3:** Correlations across frequencies

Mat.	Daily changes				Weekly changes			
	Response		Correlation		Response		Correlation	
5	0.72	(0.04)	0.46	[0.41, 0.50]	0.78	(0.09)	0.50	[0.40, 0.59]
7	0.95	(0.04)	0.58	[0.54, 0.61]	0.95	(0.09)	0.58	[0.49, 0.66]
10	0.94	(0.03)	0.65	[0.62, 0.69]	0.91	(0.07)	0.66	[0.58, 0.72]
15	0.88	(0.03)	0.61	[0.57, 0.64]	0.86	(0.07)	0.64	[0.56, 0.71]
	Monthly changes				Quarterly changes			
	Response		Correlation		Response		Correlation	
5	1.48	(0.18)	0.75	[0.61, 0.85]	1.51	(0.38)	0.70	[0.34, 0.88]
7	1.23	(0.14)	0.77	[0.63, 0.86]	1.56	(0.23)	0.86	[0.65, 0.95]
10	0.94	(0.12)	0.71	[0.55, 0.82]	1.34	(0.20)	0.85	[0.63, 0.94]
15	0.79	(0.16)	0.55	[0.33, 0.71]	1.10	(0.29)	0.66	[0.29, 0.86]

Relation between rate changes in nominal forward rates and changes in forward inflation compensation for different frequencies: response of nominal rates to IC in univariate regressions (OLS standard errors in parentheses) and correlations (with 95%-confidence intervals in squared brackets) for daily, weekly, monthly and quarterly changes.

I have shown that changes in nominal rates are highly correlated with changes in inflation compensation, and that this holds across subsamples and at various frequencies. The regression coefficients are mostly close to and often above unity. This evidence indicates that revisions of monetary policy expectations are to a large extent driven by changes in inflation expectations.

### 3.3.2 Responses to payroll surprises

An alternative approach to assess the relation between revisions to short rate expectations and revisions to inflation expectations is to consider the response to macroeconomic news. Based on the discussion in Section 3.2 and in light of the fact that expected short rates strongly respond to some macro news, such as non-farm payrolls, we would expect inflation compensation strongly respond to such news as well. However, Beechey and Wright (2009), henceforth BW, find that in intraday data inflation compensation hardly responds to real-side news. This

section first revisits the evidence of BW, zooming in on the response to payroll surprises, qualifying some of their results. I argue that we should be using daily windows to assess the response of inflation compensation to macro news, because in some cases the information gets processed more slowly. Then I provide evidence based on a new data set which shows a rather strong response of the inflation compensation to payroll surprises.

The effects of macro announcements on financial markets are typically estimated by means of an event study, where changes in asset prices around the time of a data release are regressed on a measure of the surprise component in this release (see e.g. Balduzzi et al., 2001b):

$$\Delta p_t = \beta' s_t + \varepsilon_t, \quad (3.6)$$

where  $t$  indexes days with announcements,  $\Delta p_t$  is the change in an asset price or interest rate around the announcement, and  $s_t$  is a  $k \times 1$  vector containing the surprise component for each of the  $k$  announcements – the  $k$ th element of  $s_t$  is zero if there was no release for this announcement on day  $t$ . Multivariate regression is used in order to partial out the effects of different announcements that occur on the same day. The surprise component is calculated as the difference between released number and (survey-based) expected number, standardized to have unit variance for the sake of comparability of different releases. Throughout this section the focus is on the impact of surprises in non-farm payrolls, thus I will use only those days with a new employment report, and  $s_t$  will contain the surprise component of non-farm payroll employment, the unemployment rate, and hourly earnings. The estimate of  $\beta_1$  can thus be interpreted as the impact of a one standard deviation surprise in payrolls in asset prices.

### Revisiting the Beechey-Wright evidence

Table 3.4 shows estimates of the responses (in basis points) to a surprise in non-farm payrolls for nominal rates, real (TIPS) rates, and inflation compensation. The sample corresponds exactly to the one used in BW, with the exception that only days with an employment report are used. As in BW I consider the ten-year yield, the five-year yield and the five-to-ten-year forward rate. The first two



rows show the estimated response (and White standard errors) over a 30-minute interval around the announcement. Thus the dependent variable is the change from 8:15am to 8:45am, since the release is at 8:30am. This is comparable to the results in BW's table 3, numerical differences stem from the fact that I only use the employment report data. The third and fourth row show the results for a longer intraday window, spanning four hours. The fifth and sixth row show results using daily windows as in BW's table 4 (that is the change from 4pm on the previous day to 4pm on the day of the announcement).

**Table 3.4:** Intra-daily and daily responses to non-farm payroll surprise

10y yield			5y yield			5-to-10y forward rate		
Nom.	TIPS	IC	Nom.	TIPS	IC	Nom.	TIPS	IC
<i>30-minute window</i>								
6.61	5.18	1.43	8.58	6.27	2.31	4.02	4.03	-0.01
(0.63)	(0.46)	(0.33)	(0.79)	(0.66)	(0.42)	(0.58)	(0.52)	(0.50)
<i>four-hour window</i>								
6.55	4.95	1.59	7.99	6.05	1.94	4.65	3.87	0.79
(0.91)	(0.78)	(0.24)	(1.00)	(0.92)	(0.33)	(0.88)	(0.70)	(0.26)
<i>daily window</i>								
6.70	5.03	1.66	8.23	6.25	1.98	4.70	3.76	0.94
(0.92)	(0.80)	(0.29)	(1.02)	(0.93)	(0.37)	(0.86)	(0.78)	(0.28)

Response to a one-standard-deviation surprise in non-farm payrolls of nominal rates, real rates and IC, using a 30-minute window around the announcement (8:15am to 8:45am), a 4-hour window (8:15am to 12:15pm) and a daily window (4pm previous day to 4pm on the announcement day). White standard errors in parentheses. Sample period: March-2004 to June-2008.

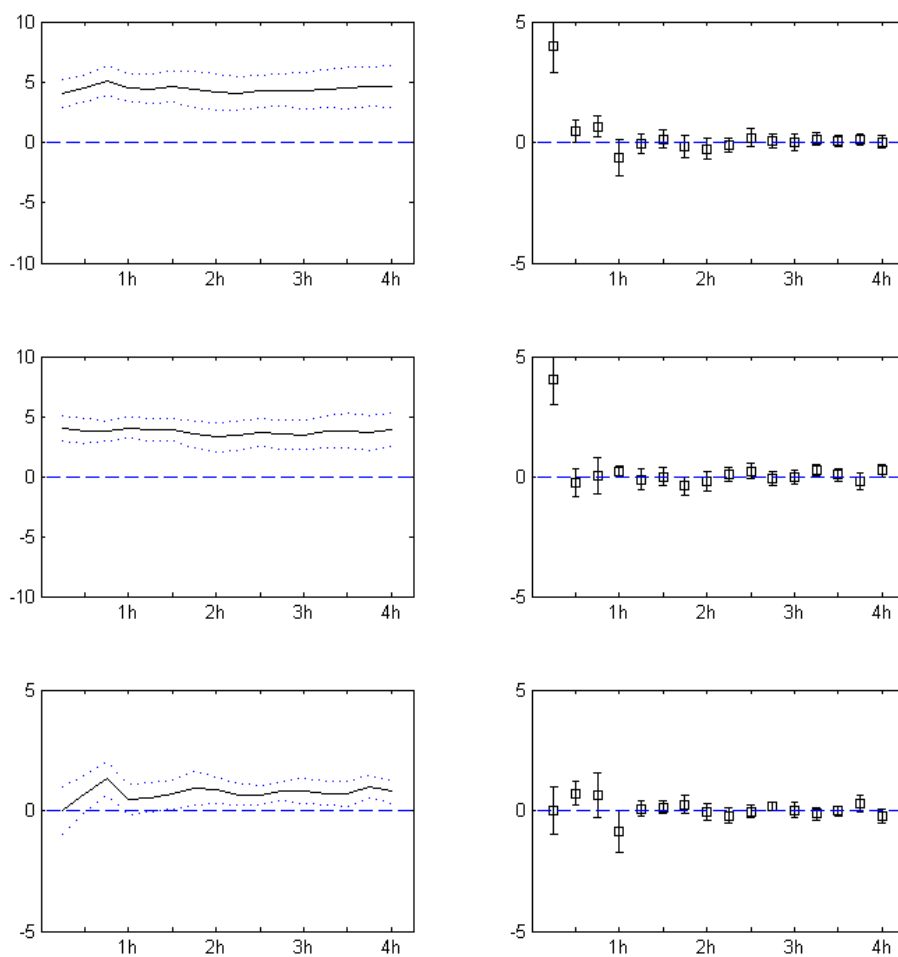
The responses of the IC to payroll surprises are rather small. Even more surprisingly, the forward IC does not respond at all to payrolls when using the 30-minute window. Generally the responses of the IC become larger for longer windows.

The striking difference between the short and long windows for the case of the forward IC warrants closer investigation of the adjustment over the course of the day. Figure 3.1 shows the response of the nominal forward rate (top row), the real forward rate (middle row), and the forward IC (bottom row) for different intra-

daily windows. The left panel shows the response coefficients and 95% confidence intervals for windows that start 15 minutes before the release and end 15 minutes to four hours after the release (12:30p) – this parallels what BW show in their Figures 2 and 3 for the ten-year nominal and real yields. It represents the cumulative response over the course of the morning. The right panel shows estimates of the responses (with 95% confidence intervals) for non-overlapping windows: the first error bar is for the 30-minute window around the release, the following ones are for the subsequent 15-minute windows until 12:30p.

While at first sight the top left and middle left panels seem to indicate that all new information is incorporated immediately into nominal and real rates and no further adjustments occur after the first 30 minutes (the conclusion of BW), this is not the whole story. The top panels in fact show that the nominal forward rate exhibits a small but significant response in the two following 15-minute intervals. This is not the case for the real forward rate (middle panels), and thus this delayed response occurs in the forward IC, as is evident in the bottom panels: The forward IC does not respond at all during the 30-minute window around the announcement, but over the subsequent periods adjusts upward.

One can only speculate as to the reasons for the delay of this response. It seems that a delayed adjustment should be arbitrated away, since such a predictable pattern would in principle generate a profitable trading opportunity: after observing a positive payroll surprise, a trader could (at 8:45am) comfortably enter into a long 5-to-10-year forward position based on nominal bonds and into a corresponding short position in TIPS bonds. This forward IC position would on average generate a positive payoff. However this is not necessarily an arbitrage opportunity for at least two reasons: First, the expected payoff might be too small to warrant the riskiness of such a trade. Second, the expected movement of about one basis point could be practically irrelevant given the prevailing bid-ask spreads in Treasury and TIPS markets – Fleming and Remolona (1999a) show that after announcements the spreads in the Treasury market significantly widen. There is evidence that even in highly liquid markets delayed adjustments of prices to macroeconomic announcements is not always arbitrated away. For example Tay-



**Figure 3.1:** Intraday responses of nominal rates, real rates and IC to payroll surprises

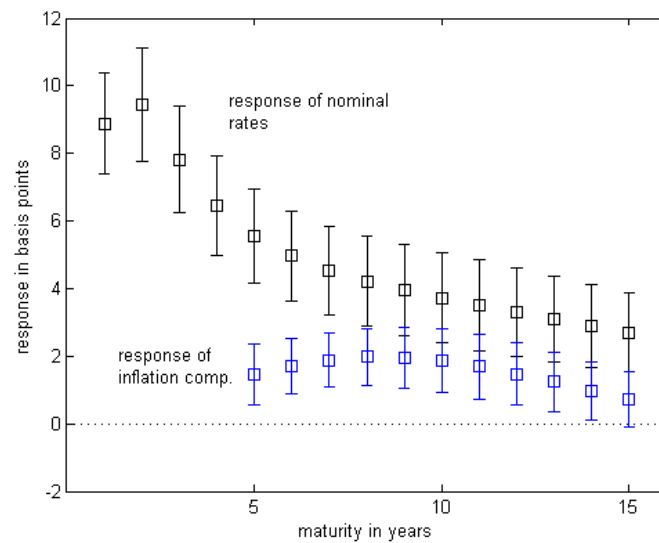
Left panel: cumulative responses to one-standard-deviation surprise in non-farm payrolls from 15 minutes before the release until 15 to 240 minutes after the release. Right panel: responses during non-overlapping windows, the first one spanning 30 minutes around the release and the following each spanning 15 minutes. Top row: nominal yield. Middle row: real yield. Bottom row: inflation compensation. Also shown are 95%-confidence intervals based on White standard errors.

lor (2010) shows that Federal funds futures adjust even until two hours after the announcement.

So should we be using tight intra-daily windows or daily changes to assess the impact of macro surprises? If the effect of macro news on asset prices is processed quickly, say within minutes, and no more processing takes place over the rest of the day, then “sizeable efficiency gains can be obtained from running these regressions with intra-daily data, rather than data at the daily frequency” (Beechey and Wright, 2009, p.536). In this case the daily change is equal to the intra-daily change (related to the surprise in the release) plus noise (unrelated to the surprise). Using daily changes would thus lead to *lower explanatory power* and *less precise estimates*, and one should instead use tight windows around the announcement. However, if for certain asset prices the information processing takes longer than several minutes, a tight intra-daily window will miss part of the announcement effect. Table 3.4 showed that the estimated response of the inflation compensation does not disappear or become insignificant with longer windows, but instead generally increases. The reason is the delayed adjustment of the forward IC to payroll surprises documented in Figure 3.1. Thus, in order to capture the full effect of payroll surprises on IC, we need to use longer windows.

### **New evidence**

Based on a larger data set including both more maturities and a longer sample period, I now present new evidence on how non-farm payroll surprises affect the nominal term structure and the term structure of inflation compensation. Daily windows are to allow for slow information processing. For the nominal term structure I use instantaneous forward rates from Gürkaynak et al. (2006b). To calculate the inflation compensation I subtract from these the instantaneous real forward rates, taken from Gürkaynak et al. (2010). Nominal rates and IC are taken to be in basis points. For both cases and for each maturity I estimate equation (3.6) separately, taking  $s_t$  to include the standardized surprises in the employment report. My sample starts in January 2003 and ends in July 2007, which amounts to 55 employment reports.



**Figure 3.2:** Responses of nominal rates and inflation compensation to non-farm payroll surprise

Response to a one standard deviation surprise in non-farm payrolls of instantaneous nominal forward rates (maturities one year to 15 years) and instantaneous forward inflation compensation (maturities five to 15 years), including 95%-confidence intervals.

Figure 3.2 shows the responses to a payroll surprise of nominal rates and IC across maturities, including 95%-confidence intervals based on OLS standard errors. Table 3.5 presents the numerical results for selected maturities, with t-statistics in parentheses. The inflation compensation shows a *significant response for essentially all maturities*. The response is hump-shaped: inflation expectations five years out respond less than those eight to ten years out, and for the longest maturities the response is smaller than for short and medium maturities.

**Table 3.5:** Responses to payroll surprises

	6y	8y	10y	12y	14y
nominal	4.96 (7.25)	4.22 (6.27)	3.73 (5.45)	3.30 (4.93)	2.90 (4.63)
IC	1.70 (4.10)	1.98 (4.56)	1.86 (3.88)	1.48 (3.10)	0.98 (2.25)
IC/nom.	34.2%	46.9%	50.0%	44.8%	33.7%

Responses to a one-standard-deviation payroll surprise of instantaneous nominal forward rates and the corresponding forward IC, with t-statistics based on OLS standard errors in parentheses. Last row: ratio of the response of the inflation compensation to the response of nominal forward rates.

To assess the importance of IC response for changes in nominal rate, consider the ratio of the two response coefficients. The last row of table 3.5 shows this ratio as a percentage number. *Between 25% and 50% of the response of nominal rates is explained by the inflation compensation.* This stands in marked contrast with the conclusion of Beechey & Wright (2009, p. 541) that “the sensitivity of nominal forward rates to news about the real side of the economy is heavily concentrated in real forward rates”. Let me clarify the reasons for the different evidence and conclusion: (i) My evidence is based on daily changes, allowing me to capture the full effect of payroll surprises on the inflation compensation, which adjusts slowly to the news. (ii) I focus on non-farm payrolls, the macro release that causes the biggest movements in nominal rates, where the difference between estimates using intra-daily and daily windows is particularly large. (iii) My sample period is different: I start my sample earlier and end it before the crisis. The financial

crisis has made some markets function less efficiently, and the liquidity premium in the TIPS market has markedly increased during that time (Lehnert et al., 2009). Excluding this extraordinary period is therefore preferable. (iv) While BW only consider two maturities – the five-year yield and the five-to-ten year forward rate (the 10-year yield is just the average of the two) – I include a broader set of maturities. This shows the role of the inflation compensation across the whole term structure.

In sum, the evidence based on TIPS data indicates that the impact of real-side macro news on inflation expectations is sizeable and contributes significantly to explaining the strong procyclical responses of monetary policy expectations.

### 3.4 Survey-Based Measures of Inflation Expectations

A natural way to measure inflation expectations is to survey market participants on a regular basis about their subjectively expected future inflation.<sup>7</sup> In this section I assess whether macroeconomic data surprises are followed by systematic changes in survey-based inflation expectations. There are three popular sources: Blue Chip Economic Indicators has conducted monthly surveys of business economists since 1976, which asks about respondents' price level expectations for each future quarter up until the end of the year following the survey. The Survey of Professional Forecasters (SPF), conducted on a quarterly basis since 1968, managed by the Federal Reserve Bank of Philadelphia since 1990, also targets business economists and asks them about their one year and ten year inflation expectations. The monthly survey of the University of Michigan's Survey Research Center has polled households since 1977 about their expectations for one year and ten year CPI inflation.

The empirical methodology is the following: Denote by  $\hat{\pi}_t^e$  the median of the reported inflation expectations in basis points in month  $t$  (for Blue Chip and

---

<sup>7</sup>Important studies using survey-based inflation expectations are, among others, Mankiw et al. (2003) and Branch (2004). For a review on this type of survey data see Lloyd (1999).

Michigan) or quarter  $t$  (for SPF) for a particular horizon (which is suppressed in the notation). Let  $s_t$  be a vector containing for each release the sum of the surprises over the month or quarter, standardized to have unit variance. The particular releases I include are non-farm payrolls, the unemployment rate, hourly earnings (the three major numbers in the employment report) as well as the Consumer Confidence Index released by the Conference Board and the CPI. These are all released once per month. Thus for the monthly frequency  $s_t$  simply contains the standardized surprise components of the releases in month  $t$ , whereas for the quarterly frequency  $s_t$  contains standardized cumulative surprises for quarter  $t$ . The following regression specification allows to estimate the responses of survey-based inflation expectations to macro news:

$$\hat{\pi}_t^e - \hat{\pi}_{t-1}^e = \alpha + \gamma_0 s_t + \gamma_1 s_{t-1} + \varepsilon_t,$$

where  $\alpha$  is an intercept, one lag of  $s_t$  is included to allow for delayed effects of macro news on survey expectations, and  $\varepsilon_t$  is an error term. The choice of the releases to be included is guided by the insight that while the surprises are orthogonal to each other, including additional relevant releases will increase efficiency by decreasing the residual variance. However I do not include all available releases since this would decrease the degrees of freedom by too much. The sample starts in January 1990 and ends in July 2007.

Results using the Michigan survey do not show any notable systematic response of inflation expectations to macro news (results not shown). Pescatori and Bianco (2009) have noted about this survey, “median forecast quite often lags actual inflation, which suggests that current inflation plays an important role in determining inflation expectations”. Thus a possible explanation for this finding is that the inflation expectations of the average consumer hardly react to macro news, in contrast to those of bond traders and business economists, because consumers pay less attention to current macro releases and base their expectations mainly on individual observations of prices and current inflation.

Focusing thus on the Blue Chip and SPF surveys, I present results for the following survey expectations: The six-quarter<sup>8</sup> Blue Chip forecast for CPI

---

<sup>8</sup>I extrapolate inflation expectations for those quarters where only four-quarter or five-quarter



inflation (BC-CPI) and for inflation based on the GDP deflator (BC-PGDP), the one year SPF forecast for inflation based on the GDP deflator (SPF-PGDP1), and the one year SPF forecast for CPI inflation (SPF-CPI1). Table 3.6 shows the estimated values for  $\gamma_0$  and  $\gamma_1$  for each survey-based expectations measure, as well as  $R^2$  and number of observations.

**Table 3.6:** Responses of survey-based inflation expectations to macro news

Lag	BC-PGDP		BC-CPI		SPF-PGDP1		SPF-CPI1	
	0	1	0	1	0	1	0	1
Payrolls	-.42 (-.84)	<b>1.24</b> (2.46)	-.35 (-.66)	-.29 (-.54)	-1.13 (-.53)	3.61 (1.65)	-3.87 (-1.64)	<b>5.43</b> (2.24)
Unempl.	-.55 (-1.11)	.22 (.44)	.27 (.52)	-.36 (-.68)	-3.41 (-1.66)	<b>-5.54</b> (-2.67)	-1.73 (-.76)	-1.49 (-.65)
Earnings	-.91 (-1.80)	-.05 (-.10)	-.32 (-.61)	-.35 (-.66)	-1.38 (-.71)	2.93 (1.51)	2.99 (1.39)	2.52 (1.18)
Cons.	.35 (.69)	<b>1.24</b> (2.44)	.51 (.96)	<b>1.18</b> (2.20)	<b>5.97</b> (2.80)	-1.89 (-.94)	<b>7.74</b> (3.28)	-.33 (-.15)
CPI	-.55 (-1.11)	<b>1.20</b> (2.44)	.12 (.24)	.57 (1.10)	-.94 (-.48)	2.41 (1.18)	2.74 (1.26)	-1.90 (-.84)
$R^2$	11.4%		4.5%		36.4%		34.4%	
$N$	211		211		70		70	

Response coefficients (t-statistics in parentheses) for survey-based expectations. Bold face indicates significance at the 5% level.

Surprises in the non-farm payroll numbers do lead to a significant increase in inflation expectations for both BC-PGDP and SPF-CPI1. However the response is delayed: not the current surprises but the lagged surprises have this effect on expectations. Notably the timing is such that the surprises are in fact observed by market participants before they provide the survey responses, so it is surprising that the response is delayed. For SPF-PGDP1 there is a barely significant delayed response.

For the BC survey all macro news have a delayed impact on expectations, if they have any significant impact. For the SPF survey the evidence is mixed with some significant contemporaneous responses and some delayed responses. One expectations are available.

problem with the SPF measures is that the number of observations is rather small (70). Given the small degrees of freedom (59) and the rather large  $R^2$  there might be a problem of over-fitting, and the SPF results should be taken with more caution.

The overall conclusion is that several survey-based measures of inflation expectations show a significant positive response to payroll surprises. Interestingly, survey-based inflation expectations do not respond as quickly as we would expect, based on the timing of the employment report data and of the survey collection.

### 3.5 Projecting Future Inflation on Macro News

The previous sections used market-based and survey-based measures of inflation expectations and estimated their relationship with macro news. An alternative is to ask how optimal forecasts of inflation should be revised in response to the new information in a data release. Since this new information is orthogonal to the information used to forecast inflation, this simply amounts to a projection of future inflation on current economic news.

Consider the monthly time series of inflation,  $\pi_t$ , as well as the surprise component in a particular monthly release, standardized to have unit variance,  $s_t$ . The object of interest is the revision to the expected future path of inflation that is caused by observing  $s_t$ ,  $E(\pi_{t+h}|s_t, \Omega_{t-1}) - E(\pi_{t+h}|\Omega_{t-1})$ , where  $\Omega_{t-1}$  is the information set available to agents before the release.<sup>9</sup> Since  $s_t$  and  $\Omega_{t-1}$  are orthogonal, the marginal revision

$$\delta^o(s_t, h) = \frac{\partial[E(\pi_{t+h}|s_t, \Omega_{t-1}) - E(\pi_{t+h}|\Omega_{t-1})]}{\partial s_t} = \frac{\partial E(\pi_{t+h}|s_t, \Omega_{t-1})}{\partial s_t},$$

is independent of  $\Omega_{t-1}$ . Here  $\delta^o(s_t, h)$  denotes the true marginal revision. If we assume that  $E(\pi_{t+h}|s_t, \Omega_{t-1})$  is linear in  $s_t$  then  $\delta^o(s_t, h) = \delta^o(h)$ . Under this assumption a linear projection of  $\pi_{t+h}$  on  $s_t$  will recover the conditional expectation.

---

<sup>9</sup>In a slight abuse of notation  $\Omega_{t-1}$  is taken to include all information that is available to agents when they form expectation of the payroll release, although of course the date of the Bloomberg survey that is used to construct  $s_t$  does not necessarily coincide with the last day of month  $t - 1$ .

Using the demeaned values for inflation, denoted by  $\tilde{\pi}_t$ , we can run the regression

$$\tilde{\pi}_{t+h} = \delta_h s_t + \varepsilon_{t,h} \quad (3.7)$$

for each horizon  $h$  and the estimates  $\hat{\delta}_h$  will be consistent for  $\delta^o(h)$ .<sup>10</sup>

The sample I use extends from March 1985 (the first month for which I have the payroll release data) to July 2007 (before the beginning of the crisis). I consider forecast horizons extending out to five years:  $h = 1, \dots, H$ , where  $H = 60$ . Thus each regression includes 209 observations, from March 1985 to July 2002, the last month for which 60-month-ahead inflation is available. As the inflation measure I use month-on-month<sup>11</sup> CPI inflation (seasonally adjusted) measured in basis points. The surprise measure  $s_t$  is taken to be the standardized surprise component in the non-farm payroll number.

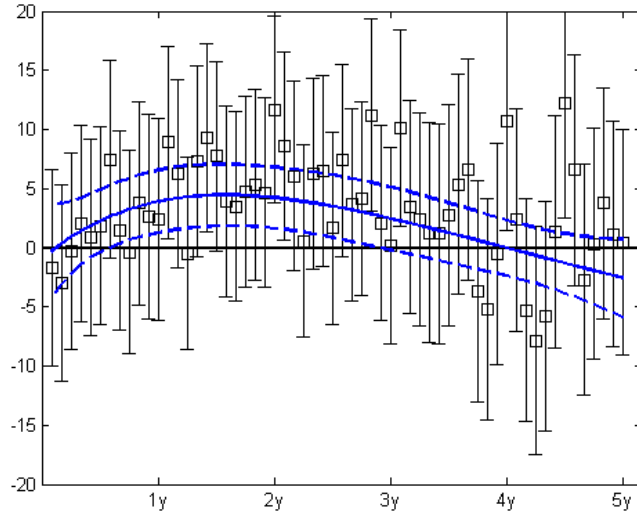
Figure 3.3 shows the estimated  $\hat{\delta}_h$  across horizons, together with 95% confidence intervals. There is a clear tendency for future inflation to be higher after a positive payroll surprise: By far most of the  $\hat{\delta}_h$  are positive, several are significantly positive, and none are significantly negative. There is a hump-shaped pattern, with forecasts for horizons of one to three years being revised upward the most.

The estimated revision of inflation forecasts caused by a payroll surprise is a high-dimensional object. The point estimates show a very jagged pattern across horizons, and the individual confidence intervals are rather large. Clearly there is a lot of noise underlying these estimates. We can improve the precision by imposing a more parsimonious structure on the revision: instead of leaving the  $\delta_h$  entirely unrestricted we can impose a smooth parametric structure. Assume that a  $k$ -dimensional parameter vector  $\theta$  determines the marginal revision:  $\delta^o(h) = \delta(\theta^o, h)$ . If we choose  $k < H$  then we have a parsimonious structure: instead of estimating  $H$  different  $\delta_h$ , each using one orthogonality condition, we estimate only  $k$  parameters which are tied down by  $H$  moment conditions

$$E[(\tilde{\pi}_{t+h} - \delta(\theta^o, h) \cdot s_t) s_t] = 0 \quad h = 1, \dots, H.$$

<sup>10</sup>Including additional predictors in (3.7) such as lagged inflation could potentially reduce the error variance and improve precision. It turns out that augmenting the regression by lagged inflation has essentially no impact on the results because forecast errors remain large.

<sup>11</sup>The use of month-on-month changes avoids the overlap that would result from the use of a monthly series of year-on-year changes.



**Figure 3.3:** Revision of optimal inflation forecast following a payroll surprise

Error bands show unrestricted projection coefficients  $\hat{\delta}_h$  for  $h = 1, \dots, 60$  including 95% confidence intervals. Solid line shows parsimoniously parameterized revision of inflation expectations following a payroll surprise,  $\delta(\hat{\theta}, h)$ . Dashed line shows point-wise 95% confidence intervals for parsimonious revision.

The parameter vector  $\theta$  is easily estimated using GMM. Let

$$h(\theta, w_t) = \begin{pmatrix} (\tilde{\pi}_{t+H} - \delta(\theta, H) \cdot s_t) s_t \\ \vdots \\ (\tilde{\pi}_{t+1} - \delta(\theta, 1) \cdot s_t) s_t \end{pmatrix}$$

so that the  $H$  moment conditions can be written as  $E[h(\theta, w_t)] = 0$ . The GMM estimator of  $\theta$  minimizes the objective function

$$Q(\theta) = \left[ T^{-1} \sum_{t=1}^T h(\theta, w_t) \right]' W \left[ T^{-1} \sum_{t=1}^T h(\theta, w_t) \right],$$

where  $W$  is a weighting matrix. If the vector process  $h(\theta^o, w_t)$  is serially uncorrelated the optimal weighting matrix is the inverse of  $S = E\{[h(\theta^o, w_t)][h(\theta^o, w_t)]'\}$ . Since  $s_t$  is a martingale difference sequence<sup>12</sup> this is indeed the case. To see this

<sup>12</sup>The definition of  $s_t$  as the surprise component in the release, i.e. the difference between the release and its conditional expectation, implies that  $E(s_t | s_{t-1}) = 0$ .

let  $\varepsilon_t = (\varepsilon_{t,1}, \dots, \varepsilon_{t,H})'$  so that  $h(\theta^o, w_t) = \varepsilon_t s_t$ , and consider

$$\begin{aligned} E[h(\theta^o, w_{t+1})|h(\theta^o, w_t)] &= E(\varepsilon_{t+1} s_{t+1} | \varepsilon_t s_t) \\ &= E[E(\varepsilon_{t+1} s_{t+1} | s_t, \varepsilon_{t+1}) | \varepsilon_t s_t] \\ &= E[\varepsilon_{t+1} E(s_{t+1} | s_t, \varepsilon_{t+1}) | \varepsilon_t s_t] \\ &= E[\varepsilon_{t+1} \cdot 0 | \varepsilon_t s_t] = 0 \end{aligned}$$

The second and third line follow from the law of iterated expectations, and the last line follows from  $E(s_{t+1} | s_t) = 0$ . To construct the optimal GMM estimator I use an iterative scheme as described in Hamilton (1994, p. 413), where first an identity weighting matrix is used to obtain an initial estimate of  $\theta$ , then  $S$  is estimated and a new estimate of  $\theta$  is obtained using  $W = \hat{S}^{-1}$ . This is repeated until the Euclidean distance between consecutive estimates of  $\theta$  is small.

As a parametric specification for the revision I use

$$\delta(\theta, h) = \theta_1 + \theta_2 e^{-\theta_4 h} + \theta_3 \theta_4 h e^{-\theta_4 h},$$

which parallels the Nelson-Siegel parametrization in the term structure literature (Diebold and Li, 2006), and is motivated by the desire to allow for a flexible hump shape while retaining parsimony.

Figure 3.3 shows the resulting parsimoniously parameterized revision of inflation expectations following a one standard deviation payroll surprise. The solid line shows the parsimoniously specified marginal revision  $\delta(\theta, h)$  evaluated at the optimal GMM estimates. The dashed line shows point-wise 95% confidence intervals, obtained using the delta method.

The key result is that in response to a payroll surprise, inflation expectations are revised significantly upward over horizons from six months to about three years. The maximum revision is about five basis points and occurs at about a two-year horizon.

What we have achieved here is that we obtained a smooth and more precisely estimated revision object. This allows much more clear cut conclusions about changes in inflation forecasts across horizons than we can make based on separate estimates for each horizon, where precision is low and a clear pattern not necessarily evident.

Since the number of moment conditions exceeds the number of parameters to be estimated, we can test whether these overidentifying restrictions are valid using Hansen's J-test (Hamilton, 1994, p. 415). The value of the test statistic is 51.85, while the 5% critical value of a  $\chi^2$ -distribution with  $H - \dim(\theta) = 60 - 4 = 56$  degrees of freedom is 74.47. Thus we do not reject the parametric model chosen to restrict the shape of the revision.

In sum, based on linear projections of future realizations of inflation on current payroll surprises, inflation forecasts should be significantly revised upwards following a positive surprise in the payroll numbers. The magnitude of these revision is quite sizeable at horizons of about two years. This is in line with the evidence from the previous sections, confirming that changing inflation expectations likely play an important role in the strong pro-cyclical response of nominal rates to real-side macro news.

### 3.6 Conclusion

In this paper I have shown that revisions to inflation expectations are an important source for variation in nominal short rate expectations, i.e. for revisions of the expected path of future monetary policy. Market-based measures of revisions of inflation expectations and of policy expectations show strong correlations, independent of the particular sub-sample or frequency considered. Furthermore I showed that real-side macro news, such as surprises in non-farm payroll employment, cause significant revisions in inflation expectations. Alternative ways to measure changes in inflation expectations, based on market data, survey data and optimal forecasts, all lead to this conclusion.

There are some interesting possible extensions to this work. Market-based measures of inflation expectations, based on TIPS and nominal yields, can be improved by taking into account variation in inflation risk premia and liquidity premia (see for example D'Amico et al., 2010). With regard to the survey-based measures, it seems desirable to combine the different available survey estimates, possibly incorporating additional sources, to obtain one single survey-based mea-

sure of the expected path of inflation expectation. And for the estimates based on future realizations of inflation, a successful forecast model could improve the precision of the estimates by reducing the projection error variance. While these extensions are unlikely to alter the rather robust empirical conclusion of this paper, they might well improve the quality of the estimates. In particular these extensions might help to improve the quantitative consistency of the results of the different approaches, in terms of the magnitude and the shape of the revisions of inflation expectations that are caused by particular macro news.

# Bibliography

- Ang, Andrew and Monika Piazzesi**, “A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables,” *Journal of Monetary Economics*, May 2003, 50 (4), 745–787.
- , **Sen Dong, and Monika Piazzesi**, “No-Arbitrage Taylor Rules,” NBER Working Papers 13448, National Bureau of Economic Research, Inc September 2007.
- Backus, David K., Chris I. Telmer, and Liuren Wu**, “Design and Estimation of Affine Yield Models,” GSIA Working Papers 5, Carnegie Mellon University, Tepper School of Business November 1999.
- Balduzzi, Pierluigi, E. J. Elton, and T. C. Green**, “Economic News and Bond Prices: Evidence from the U.S. Treasury Market,” *Journal of Financial and Quantitative Analysis*, December 2001, 36 (4), 523–544.
- , – , **and** – , “Economic News and Bond Prices: Evidence from the U.S. Treasury Market,” *Journal of Financial and Quantitative Analysis*, December 2001, 36 (4), 523–544.
- , **Sanjiv Ranjan Das, and Silverio Foresi**, “The Central Tendency: A Second Factor In Bond Yields,” *The Review of Economics and Statistics*, February 1998, 80 (1), 62–72.
- , – , – , **and Rangarajan Sundaram**, “A Simple Approach to Three-Factor Affine Term Structure Models,” *Journal of Fixed Income*, December 1996, pp. 43–53.



- Bansal, Ravi and Hao Zhou**, “Term Structure of Interest Rates with Regime Shifts,” *Journal of Finance*, October 2002, 57 (5), 1997–2043.
- Bartlett, M. S.**, “A comment on D. V. Lindley’s statistical paradox,” *Biometrika*, 1957, 44 (3-4), 533–534.
- Beechey, Meredith**, “A closer look at the sensitivity puzzle: the sensitivity of expected future short rates and term premia to macroeconomic news,” Finance and Economics Discussion Series 2007-06, Board of Governors of the Federal Reserve System (U.S.) 2007.
- Beechey, Meredith J. and Jonathan H. Wright**, “The high-frequency impact of news on long-term yields and forward rates: Is it real?,” *Journal of Monetary Economics*, May 2009, 56 (4), 535–544.
- , **Benjamin K. Johansson**, , and **Andrew T. Levin**, “Are Long-Run Inflation Expectations Anchored More Firmly in the Euro Area than in the United States?,” Finance and Economics Discussion Series Working Paper, Federal Reserve Board of Governors September 2007.
- Bertholon, H., A. Monfort, and F. Pegoraro**, “Econometric Asset Price Modelling,” *Journal of Financial Econometrics*, 2008, 6 (4), 407–458.
- Boivin, Jean, Sen Dong, and Andrew Ang**, “Monetary Policy Shifts and the Term Structure,” *AFA 2009 San Francisco Meetings Paper*, March 2009.
- Branch, William A.**, “The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations,” *The Economic Journal*, 2004, 114 (497), 592–621.
- Brockwell, Peter J. and Richard A. Davis**, *Time Series: Theory and Methods*, second ed., Springer, 2006.
- Brooks, Stephen P. and Gareth O. Roberts**, “Convergence assessment techniques for Markov chain Monte Carlo,” *Statistics and Computing*, Dec 1998, 8 (4), 319–335.

- Campbell, John Y., Adi Sunderam, and Luis M. Viceira**, “Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds,” NBER Working Papers 14701, National Bureau of Economic Research, Inc Feb 2009.
- Carlin, Bradley P. and Siddhartha Chib**, “Bayesian Model Choice via Markov Chain Monte Carlo Methods,” *Journal of the Royal Statistical Society. Series B (Methodological)*, September 1995, 57 (3), 473–484.
- Carter, C. K. and R. Kohn**, “On Gibbs sampling for state space models,” *Biometrika*, 1994, 81 (3), 541–553.
- Chib, Siddhartha**, “Markov chain Monte Carlo methods: computation and inference,” in J.J. Heckman and E.E. Leamer, eds., *Handbook of Econometrics*, Vol. 5, Elsevier, 2001, chapter 57, pp. 3569–3649.
- **and Edward Greenberg**, “Understanding the Metropolis-Hastings Algorithm,” *The American Statistician*, Nov 1995, 49 (4), 327–335.
- **and –**, “Markov Chain Monte Carlo Simulation Methods in Econometrics,” *Econometric Theory*, Aug 1996, 12 (3), 409–431.
- Christensen, Jens H.E., Francis X. Diebold, and Glenn D. Rudebusch**, “The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models,” Working Paper Series 2007-20, Federal Reserve Bank of San Francisco 2007.
- Clarida, Richard, Jordi Gali, and Mark Gertler**, “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory,” *The Quarterly Journal of Economics*, 2000, 115 (1), 147–180.
- Clyde, Merlise and Edward I. George**, “Model Uncertainty,” *Statistical Science*, 2004, 19 (1), 81–94.
- Cochrane, John**, “Commentary on “Macroeconomic implications of changes in the term premium”,” *Federal Reserve Bank of St. Louis Review*, Jul 2007, pp. 271–282.

- Cochrane, John H.**, “Identification with Taylor Rules: A Critical Review,” NBER Working Papers 13410, National Bureau of Economic Research, Inc September 2007.
- **and Monika Piazzesi**, “Bond Risk Premia,” *American Economic Review*, March 2005, *95* (1), 138–160.
- **and –**, “Decomposing the Yield Curve,” manuscript 2008.
- Cowles, Mary Kathryn and Bradley P. Carlin**, “Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review,” *Journal of the American Statistical Association*, 1996, *91*, 883–904.
- Dai, Qiang and Kenneth J. Singleton**, “Specification Analysis of Affine Term Structure Models,” *The Journal of Finance*, Oct 2000, *55* (5), 1943–1978.
- **and –**, “Expectation puzzles, time-varying risk premia, and affine models of the term structure,” *The Journal of Financial Economics*, Mar 2002, *63* (3), 415–441.
- **and Kenneth Singleton**, “Term Structure Dynamics in Theory and Reality,” *Review of Financial Studies*, 2003, *16* (3), 631–678.
- **, Anh Le, and Kenneth J. Singleton**, “Discrete-time Dynamic Term Structure Models with Generalized Market Prices of Risk,” working paper Mar 2006.
- D’Amico, Stefania, Don H. Kim, and Min Wei**, “Tips from TIPS: the informational content of Treasury Inflation-Protected Security prices,” Finance and Economics Discussion Series 2010-19, Board of Governors of the Federal Reserve System (U.S.) 2010.
- Del Negro, Marco, Gauti Eggertsson, Andrea Ferrero, and Nobuhiro Kiyotaki**, “The Great Escape? A Quantitative Evaluation of the Fed’s Non-Standard Policies,” working paper, Federal Reserve Bank of New York and Princeton University March 2010.

- Dellaportas, Petros, Jonathan J. Forster, and Ioannis Ntzoufras**, “On Bayesian model and variable selection using MCMC,” *Statistics and Computing*, January 2002, *12* (1), 27–36.
- DiCiccio, Thomas J., Robert E. Kass, Adrian Raftery, and Larry Wasserman**, “Computing Bayes Factors by Combining Simulation and Asymptotic Approximations,” *Journal of the American Statistical Association*, Sep 1997, *92* (439), 903–915.
- Diebold, Francis X. and Canlin Li**, “Forecasting the term structure of government bond yields,” *Journal of Econometrics*, 2006, *130*, 337–364.
- Duffee, Gregory R.**, “Term Premia and Interest Rate Forecasts in Affine Models,” *Journal of Finance*, 02 2002, *57* (1), 405–443.
- , “Forecasting with the Term Structure: the Role of No-Arbitrage,” Working Paper June, Haas School of Business, University of California, Berkeley 2009.
- **and Richard H. Stanton**, “Estimation of Dynamic Term Structure Models,” working paper, Haas School of Business Mar 2004.
- **and –** , “Evidence on Simulation Inference for Near Unit-Root Processes with Implications for Term Structure Estimation,” *Journal of Financial Econometrics*, Winter 2008, *6* (1), 108–142.
- Duffie, Darrell and Rui Kan**, “A Yield-Factor Model of Interest Rates,” *Mathematical Finance*, 1996, *6* (4), 379–406.
- Fama, Eugene F. and Robert R. Bliss**, “The Information in Long-Maturity Forward Rates,” *The American Economic Review*, 1987, *77* (4), 680–692.
- Faust, Jon, John H. Rogers, Shing-Yi B. Wang, and Jonathan H. Wright**, “The high-frequency response of exchange rates and interest rates to macroeconomic announcements,” *Journal of Monetary Economics*, May 2007, *54* (4), 1051–1068.

- Fleming, Michael J. and Eli M. Remolona**, “What moves the bond market?,” *Economic Policy Review*, December 1997, pp. 31–50.
- and –, “Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information,” *Journal of Finance*, October 1999, *54* (5), 1901–1915.
- and **Eli M Remolona**, “The term structure of announcement effects,” BIS Working Papers 71, Bank for International Settlements June 1999.
- Gamerman, Dani and Hedibert F. Lopes**, *Markov Chain Monte Carlo*, second ed., Chapman & Hall/CRC, 2006.
- Gelman, Andrew and Donald B. Rubin**, “Inference from Iterative Simulation Using Multiple Sequences,” *Statistical Science*, 1992, *7* (4), 457–472.
- George, Edward I. and Robert E. McCulloch**, “Variable Selection Via Gibbs Sampling,” *Journal of the American Statistical Association*, September 1993, *88* (423), 881–889.
- Geweke, John**, “Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments,” in “In Bayesian Statistics 4” Oxford University Press 1992, pp. 169–193.
- Green, Peter J.**, “Reversible jump Markov chain Monte Carlo computation and Bayesian model determination,” *Biometrika*, 1995, *82* (4), 711–732.
- Gürkaynak, Refet S., Andrew T. Levin, and Eric T. Swanson**, “Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from Long-Term Bond Yields in the U.S., U.K., and Sweden,” Working Paper, Federal Reserve Bank of San Francisco March 2006.
- , **Brian P. Sack, and Eric T. Swanson**, “Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements,” *International Journal of Central Banking*, May 2005, *1* (1), 55–93.

- , – , and – , “The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models,” *American Economic Review*, March 2005, *95* (1), 425–436.
- , – , and **Jonathan H. Wright**, “The TIPS Yield Curve and Inflation Compensation,” *American Economic Journal: Macroeconomics*, 2010, *2* (1), 70–92.
- , **Brian Sack**, and **Jonathan H. Wright**, “The U.S. Treasury yield curve: 1961 to the present,” Finance and Economics Discussion Series 2006-28, Board of Governors of the Federal Reserve System (U.S.) 2006.
- Hamilton, James D.**, *Time Series Analysis*, Princeton University Press, 1994.
- , “Daily Changes in Fed Funds Futures Prices,” NBER Working Papers 13112, National Bureau of Economic Research, Inc May 2007.
- , “Assessing monetary policy effects using daily federal funds futures contracts,” *Federal Reserve Bank of St. Louis Review*, Jul 2008, pp. 377–394.
- , **Seth Pruitt**, and **Scott C. Borger**, “The Market-Perceived Monetary Policy Rule,” working paper, University of California, San Diego Aug 2009.
- Hansen, Bruce E.**, “Averaging Estimators for Autoregressions with a Near Unit Root,” *Journal of Econometrics*, forthcoming 2009.
- Hoeting, Jennifer A.**, **David Madigan**, **Adrian E. Raftery**, and **Chris T. Volinsky**, “Bayesian Model Averaging: A Tutorial,” *Statistical Science*, Nov 1999, *14* (4), 382–401.
- Jardet, Caroline**, **Alain Monfort**, and **Fulvio Pegoraro**, “No-arbitrage near-cointegrated VAR(p) term structure models, term premia and GDP growth,” working paper, Banque de France Jun 2009.
- Johannes, Michael** and **Nicholas Polson**, “MCMC Methods for Continuous-Time Financial Econometric,” in Yacine Ait-Sahalia and Lars Hansen, eds., *Handbook of Financial Econometrics*, Vol. 2, Elsevier, 2009, chapter 1.

- Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu**, “A New Perspective on Gaussian DTSMs,” working paper Jul 2009.
- , **Marcel Pribsch, and Kenneth J. Singleton**, “Risk Premium Accounting in Macro-Dynamic Term Structure Models,” working paper Nov 2008.
- Kass, Robert E. and Adrian E. Raftery**, “Bayes Factors,” *Journal of the American Statistical Association*, jun 1995, *90* (430), 81–94.
- Kendall, M. G.**, “A note on bias in the estimation of autocorrelation,” *Biometrika*, 1954, *41* (3-4), 403–404.
- Kim, Chang-Jin and Charles R. Nelson**, *State-Space Models with Regime-Switching*, The MIT Press, 1999.
- Kim, Don H.**, “Challenges in macro-finance modeling,” BIS Working Papers 240, Bank for International Settlements December 2007.
- **and Athanasios Orphanides**, “Term structure estimation with survey data on interest rate forecasts,” Finance and Economics Discussion Series 2005-48, Board of Governors of the Federal Reserve System (U.S.) 2005.
- **and** – , “The bond market term premium: what is it, and how can we measure it?,” *BIS Quarterly Review*, June 2007.
- **and Jonathan H. Wright**, “An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates,” Finance and Economics Discussion Series 2005-33, Board of Governors of the Federal Reserve System (U.S.) 2005.
- King, Robert G.**, “The new IS-LM model: language, logic, and limits,” *Economic Quarterly*, Summer 2000, *86* (3), 45–103.
- Kuttner, Kenneth N.**, “Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market,” *Journal of Monetary Economics*, 2001, *47*, 523–544.

- Lehnert, Thorsten, Aleksandar Andonov, and Florian Bardong**, “TIPS, Inflation Expectations and the Financial Crisis,” LSF Research Working Paper Series 09-09, Luxembourg School of Finance, University of Luxembourg 2009.
- Litterman, Robert and J. Scheinkman**, “Common Factors Affecting Bond Returns,” *The Journal of Fixed Income*, June 1991, 1 (1), 54–61.
- Lloyd, B. Thomas Jr.**, “Survey Measures of Expected U.S. Inflation,” *The Journal of Economic Perspectives*, 1999, 13 (4), 125–144.
- Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers**, “Disagreement about Inflation Expectations,” *NBER Macroeconomics Annual*, 2003, 18, 209–248.
- Marmol, Francesc**, “The Stationarity Conditions for an AR(2) Process and Schur’s Theorem,” *Econometric Theory*, December 1995, 11 (5), 1180–1182.
- Monfort, Alain and Fulvio Pegoraro**, “Switching VARMA Term Structure Models,” *Journal of Financial Econometrics*, 2007, 5 (1), 105–153.
- Nelson, Charles R. and Andrew F. Siegel**, “Parsimonious Modeling of Yield Curves,” *The Journal of Business*, 1987, 60 (4), 473–489.
- Orphanides, Athanasios**, “Taylor rules,” in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan, 2008.
- Peres-Neto, Pedro R., Donald A. Jackson, and Keith M. Somers**, “How many principal components? stopping rules for determining the number of non-trivial axes revisited,” *Computational Statistics & Data Analysis*, June 2005, 49 (4), 974–997.
- Pescatori, Andrea and Timothy Bianco**, “Inflation and Inflation Expectations,” *Economic Trends*, May 2009.



- Piazzesi, Monika**, “An Econometric Model of the Yield Curve with Macroeconomic Jump Effects,” NBER Working Papers 8246, National Bureau of Economic Research, Inc April 2001.
- , “Bond Yields and the Federal Reserve,” *Journal of Political Economy*, April 2005, *113* (2), 311–344.
- , “Affine Term Structure Models,” in Yacine Ait-Sahalia and Lars Hansen, eds., *Handbook of Financial Econometrics*, Vol. 1, Elsevier, 2009, chapter 12.
- **and Eric T. Swanson**, “Futures prices as risk-adjusted forecasts of monetary policy,” *Journal of Monetary Economics*, May 2008, *55* (4), 677–691.
- Poole, William and Robert H. Rasche**, “Perfecting the Market’s Knowledge of Monetary Policy,” *Journal of Financial Services Research*, 2000, *18* (2/3), 255–298.
- Raftery, Adrian E. and Steven Lewis**, “How Many Iterations in the Gibbs Sampler?,” in “In Bayesian Statistics 4” Oxford University Press 1992, pp. 763–773.
- Rigobon, Roberto and Brian Sack**, “The impact of monetary policy on asset prices,” *Journal of Monetary Economics*, November 2004, *51* (8), 1553–1575.
- **and** – , “Noisy Macroeconomic Announcements, Monetary Policy, and Asset Prices,” NBER Working Papers 12420, National Bureau of Economic Research, Inc August 2006.
- Rose, Andrew Kenan**, “Is the Real Interest Rate Stable?,” *Journal of Finance*, December 1988, *43* (5), 1095–1112.
- Rotemberg, Julio J. and Michael Woodford**, “Interest Rate Rules in an Estimated Sticky Price Model,” in “Monetary Policy Rules” NBER Chapters, National Bureau of Economic Research, Inc, June 1999, pp. 57–126.
- Rudebusch, Glenn D.**, “Monetary Policy Inertia: Fact or Fiction?,” *International Journal of Central Banking*, December 2006, *2* (4).

- **and Tao Wu**, “A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy,” *Economic Journal*, 07 2008, 118 (530), 906–926.
- , **Eric T. Swanson, and Tao Wu**, “The Bond Yield “Conundrum” from a Macro-Finance Perspective,” *Monetary and Economic Studies*, December 2006, 24 (S1), 83–109.
- Startz, Richard and Kwok Ping Tsang**, “The Yield Curve through Time and Across Maturities,” Working Papers UWEC-2007-05, University of Washington, Department of Economics March 2007.
- Swanson, Eric T.**, “What we do and don’t know about the term premium,” *FRBSF Economic Letter*, Jul 2007, 20.
- Taylor, John B.**, “Discretion versus policy rules in practice,” *Carnegie-Rochester Conference Series on Public Policy*, 1993, 39, 195–214.
- , “A Historical Analysis of Monetary Policy Rules,” in “Monetary Policy Rules” NBER Chapters, National Bureau of Economic Research, Inc, June 1999, pp. 319–348.
- Taylor, Nicholas**, “The Determinants of Future U.S. Monetary Policy: High-Frequency Evidence,” *Journal of Money, Credit and Banking*, March 2010, 42 (2-3), 399–420.
- Timmermann, Allan**, “Forecast Combinations,” in G. Elliott, C. Granger, and A. Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 1, Elsevier, 2006, chapter 4.
- White, Halbert**, “Maximum Likelihood Estimation of Misspecified Models,” *Econometrica*, January 1982, 50 (1), 1–25.
- Zellner, Arnold**, “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias,” *Journal of the American Statistical Association*, June 1962, 57 (298), 348–368.