A Dynamic Term Structure Model of Central Bank Policy

by

Shawn W. Staker

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2009

Author .....
Department of Electrical Engineering and Computer Science
June5, 2009

Certified by....
Leonid Kogan
Nippon Telephone and Telegraph Professor of Management
Thesis Supervisor

Accepted by .....
Terry P. Orlando
Chairman, Department Committee on Graduate Theses

A Dynamic Term Structure Model of Central Bank Policy

by

Shawn W. Staker

Submitted to the Department of Electrical Engineering and Computer Science
on June 5, 2009, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis investigates the implications of explicitly modeling the monetary policy of the Central Bank within a Dynamic Term Structure Model (DTSM). We follow Piazzesi (2005) and implement monetary policy by including the Fed target rate as a state variable. The discontinuous target dynamics are accurately modeled via a non-linear switching process, while still maintaining affine requirements under the pricing measure ensuring tractability. To ensure a flexible risk specification we turn to the parametrization of Cheridito et al (2007), with extensions to the target jump process. Model parameters are estimated via a simulated maximum likelihood estimation scheme with importance sampling. A Bayesian particle filter is used as a robustness check, and it's use for static parameter estimation in a DTSM framework is explored.

Our results support those in Piazzesi (2005), revealing a substantial improvement in pricing errors especially on the short end of the yield curve. The model construction provides a natural framework to inspect monetary policy information embedded in yields, which is found to be substantial. We find the addition of the target rate greatly improves the model's ability to explain excess return. An ability which is increased with the inclusion of the full term structure of target rates, as measured from Fed future contracts. We postulate the improved performance is due to the target as a proxy for short term rates, and a conduit to express the information content of the term structure of target rates.

Thesis Supervisor: Leonid Kogan
Title: Nippon Telephone and Telegraph Professor of Management

Acknowledgments

I would like to thank my research advisor Leonid Kogan for his continual support and never ending patience. I am also grateful to the members of my research committee: Munther Dahleh, Scott Joslin, and John Tsitsiklis. Much of my work, and most of my sanity, is due to the invaluable and extensive discussions with Scott Joslin. Teaching for Munther Dahleh and John Tsitsiklis are highlights of my MIT education, providing experiences which have shown me *my* way forward.

I owe a special debt to Andrew Lo for providing me with office space at the Laboratory for Financial Engineering. A unique collaborative center, supporting a wide range of bright students and exciting research. Finally if it wasn't for fellow classmate Amir Khandani, I would have driven myself mad with talk of Q measures.

Though the trials and tribulations associated with this thesis have marked my time at MIT, the meeting of Tufool Al-Nuaimi has marked my life. I owe her more than I can ever repay, and love her more than I can say. My academic accomplishments pail in comparison to the pride I feel in starting a new chapter of my life with the woman I love. Ahibik b'kul qalbee.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis investigates the implications of explicitly modeling the monetary policy of the Central Bank within a Dynamic Term Structure Model (DTSM). We follow Piazzesi (2005) and implement monetary policy by including the Fed target rate as a state variable. The discontinuous target dynamics are accurately modeled via a non-linear switching process, while still maintaining affine requirements under the pricing measure ensuring tractability. To ensure a flexible risk specification we turn to the parametrization of Cheridito et al (2007), with extensions to the target jump process. Model parameters are estimated via a simulated maximum likelihood estimation scheme with importance sampling. A Bayesian particle filter is used as a robustness check, and it's use for static parameter estimation in a DTSM framework is explored.

Our results support those in Piazzesi (2005), revealing a substantial improvement in pricing errors especially on the short end of the yield curve. The model construction provides a natural framework to inspect monetary policy information embedded in yields, which is found to be substantial. We find the addition of the target rate greatly improves the model's ability to explain excess return. An ability which is increased with the inclusion of the full term structure of target rates, as measured from Fed future contracts. We postulate the improved performance is due to the target as a proxy for short term rates, and a conduit to express the information content of the term structure of target rates.

Motivation

Our motivation to explicitly include monetary policy into a term structure model stems from it's well established importance in the wider economy. Changes in the monetary base directly effect short term rates, and though various intermediaries effect everything from consumer credit to exchange rates. The Fed alters the monetary base to meet it's dual mandate of price stability and sustainable economic growth [36]. The Federal Reserve Bank maintains three controls for monetary policy, all with the purpose of influencing the daily Fed funds rate. The Fed funds rate is the overnight interest rate which depository institutions lend balances to other depository institutions. Based on this description a logical choose would be to use the Fed funds rate as a state variable reflecting monetary policy. Unfortunately the realized Funds rate has a number of adverse qualities, principally very high volatility which is found in many money market rates. As with many money market rates, spikes in the rates often appear due to known institutional constraints rather than economic drivers.

An alternative choice is to use the publicly announced Fed funds target rate, or target rate. This is the rate the Fed attempts to steer the funds rate toward via it's tools of monetary policy. The process in which the Fed conveys the target rate has evolved with the organization's view on transparency. In 1994 the Federal Open Market Committee (FOMC) began disclosing changes to it's policy stance. This evolved in 1995 to a full announcement of the current target level. Communication tools have evolved since then in an effort to increase transparency, notably in 2000 when the FOMC began to issue assessments on risks to it's dual mandate[1]. Due in part to data constraints, our analysis will be limited to a post 1995 time period where changes to the target rate are publicly announced. Besides the lack of volatility found in the funds rate, the target is appealing for two broad reasons. The first is the FOMC's ability to keep the funds rate close to the target rate, which is discussed in more detail in section 3.2. The second is the unique trait in that the target is not a market rate, thus it contains no risk premium. As such it may be seen as a proxy for other indicators of the economy, such as inflation or GDP. This fact will have

[1] A full description from the Fed's perspective is available at www.federalreserve.gov

interesting implications for risk premium as discussed in section 4.4.

From a mathematical modeling perspective, the target possesses several unique characteristics with respect to other interest rates. Since 1994 the FOMC has adjusted the target rate in 25 basis point (bp) increments, and almost predominately at one of the eight annual FOMC meetings. Incorporating the unique discretized dynamics and the strong seasonality has the potential to greatly increase the model's ability to explain observed phenomenon.

From an econometric perspective there have been a number of recent studies lending support to the importance of the Fed target. One of first studies on yield responses to changes in the target showed mixed results [16]. Once target changes were decomposed into expected and unexpected changes studies found significant yield response even with long dated maturities [45]. That expected target changes should have no effect on yields, makes intuitive sense as this information has already been incorporated into prices. The fact that unexpected target changes, or shocks, effect yields is a strong motivator for our study. Event studies also provide strong support of jumps during FOMC announcements, such as [34]. Finally more parametric models have provided strong support in allowing interest rates to jump during FOMC announcements, see [43] and [42].

Literature Review

The reported work follows several themes in the current literature. Most notably are other published studies which explicitly model the Fed target rate within a term structure model. The most notable of which are [40], [50], and [51]. [40] investigate a Gaussian term structure model with Fed targeting, where deterministic Fed jumps are used. [50] build a finite state Markov Chain to characterize target changes, and find good predictability over a somewhat short sample period.

In particular our work can be seen as an extension of the model in [51], who uses an affine term structure model with conditionally Poisson jumps driving the target. Our model uses a richer specification for the jump process, combining a single time series with a non-linear switching mechanism to drive target jumps. Other

extensions include a more flexible parametric description of the pricing kernel, and an estimation scheme incorporating a variance reduction technique. We also exploit a Bayesian Particle filter for a robustness check, and explore it's use for parameter estimation. Finally our results focus more heavily on model risk premium, and it's response to the market's term structure of target rates.

Our modeling approach falls within the broad class of Dynamic Term Structure models (DTSM). To maintain tractability we transform the dynamics under the pricing measure which places the model under the affine variation of DTSM models, or Affine Term Structure Models (ATSM). [59] and [17] are early examples of term structcure models, which are now seen as specific examples of ATSMs. The seminal work which established the broad framework of ATSM models is found in [28] or [27]. Since then [19] and [20] further established the foundations of drift-diffusion ATSM models which exist today. This framework was extended to in [29] to accomidates Poisson type jumps. An excellent survey paper on ATSM models is found in [52], with an equivilant in the DTSM space available in [18]. Finally a comprehensive text book treatment of DTSM models with an empirical focus is found in [58].

Within ATSM modoels the pursuit of more flexible parametric specifications for risk premium has been an active subject. An appropriate starting point is the completely affine price of risk specification in [19]. A slight adjustment in [23] was later coined semi-affine, followed by essentially affine in [24]. These variations have led to the fully flexible extended affine configuration of [12], which allows for a price of risk such that the state space under P and Q are fully flexible affine constructions. This is the method we apply in section 3.2.3, and later extend for jumps.

A key ingredient to any term structure model is the means to calibrate static model parameters to historical data, and possibly infer latent state variables. We will label all such issues under the umbrella of estimation. We identify the two work horses of ATSM estimation as Method of Moments and Maximum Likelihood. A comprehensive summary of each method applied to DTSM models is available in [58]. As the more efficient estimator we focus our efforts on maximum likelihood techniques. The likelihood function associated with our model dynamics is not known

in closed form and is computationally expensive to construct. This leads us to follow the existing vein of literature in simulated maximum likelihood (SML) techniques. This Monte Carlo technique was first presented for drift-diffusion models by [49]. An excellent empirical comparison of SML techniques with variance reduction techniques is available in [30]. To reduce the variance in our technique we turn to the importance sampler proposed in [35], and later formalized in [32].

A significant percentage of ATSM models include latent state variables. Within the literature a popular means to infer an N dimensional latent state space is to assume N measurements are observed without error [58]. An alternative is to use a filtering technique to infer latent states and compute the likelihood function. For Gaussian systems the infamous Kalman filter is the ideal choice, and a popular approximation when normality is not present [47]. An alternative to the Kalman filter for non-linear non-Gaussian systems is the Bayesian Particle Filter. Excellent summary papers from an engineering perspective are [2, 11], while the technique is presented for financial problems in [44]. Though flexible and robust, particle filters suffer from high variance resulting in a discontinuous likelihood functions. A discontinuous likelihood function with respect to parameters restricts the use of any gradient based optimization routines. This trait has greatly limited it's use in static parameter estimation. Considering this constraint we leverage the filter construction in [54] to quantify the robustness of our estimates and the ensuing results.

Outline of Thesis

Chapter one contains a detailed presentation on Dynamic and Affine Term Structure Models, providing much of the background for the following chapters. Chapter two presents our model construction and further estimation details. The model construction includes additional details on the Federal Reserve Bank and the model incorporates monetary policy. Chapter three contains a full discussion of results, with emphasis on pricing error and risk premium. We state our conclusions in chapter four.

Chapter 2

Dynamic Term Structure Models

Dynamic term structure models (DTSM) are mathematical models which ensure consistent joint evolution of the yield curve through time. Relative to other modeling techniques DTSMs provide a consistent framework for cross sectional pricing, excluding prices which allow arbitrage. They also possess well defined dynamics in time, allowing characterization of historical changes in the yield curve. Since DTSMs provide a complete probabilistic model for the yield curve, prices for any fixed income security may be constructed including derivatives.

In this chapter we provide a brief overview of Dynamic Term Structure Models. Section 2.1 presents a review of the mathematical foundation from which DTSM models are based. Section 2.2 introduces the affine class of DTSM models, affine term structure models (ATSM). To achieve tractability we constrain our model dynamics under the pricing measure, thus placing the model within the ATSM framework. Section 2.2 contains background discussions regarding a number of key ingredients for any ATSM model.

2.1 Mathematical Foundation

Assume the existence of a scalar instantaneous interest rate, $r(t)$, which can be written as a function of an Markov process, $X_t \in \mathbb{R}^N$ defined on a probability space $(\Omega, \mathcal{F}, P)$.

That is

$$r(t) \rightarrow r(X_t, t) \tag{2.1}$$

Define a bond as a contract which pays one dollar at time T, and has a time t price $B(t, T, X_t)$ given by

$$B(t, T, X_t) = E_t^Q \left[\exp\left(- \int_t^T r(X_s, s) ds \right) \middle| F_t \right] \tag{2.2}$$

Where the expectation is taken over a measure Q, equivalent to P, and F_t denotes a filtration with respect to time t. Note the functional dependence of B on X_t is a direct result of the Markovian assumption. Heuristically, the reason to change measure is to compensate investors for the risk they bear, where X_t may be viewed as risk factors. Existence of Q and a solution to eq(2.2) is assured under assumptions of no arbitrage, with regulatory conditions on r and dX_t. The discussion of measures is relegated to section 2.2.

To solve for model implied bond prices, one must solve the conditional expectation in eq(2.2). Possible methods include (i) direct evaluation of the conditional expectation in closed form, (ii) numerical approximation via a Monte Carlo technique, or (iii) mapping the conditional expectation to a partial differential equation (PDE) via Feynman-Kac. Observed prices strongly reject the use of X_t dynamics required for a closed form solution. Though Monte Carlo is a convenient method to generate prices, it is computationally prohibitive when calibrating the model. Furthermore Monte Carlo may be computationally infeasible when X_t is latent, a common trait in DTSMs. For these reasons the vast majority of research exploits the *Feynman-Kac stochastic representation formula* to construct a solution to eq(2.2).

Consider a PDE for $T > 0$ of the form

$$\begin{gathered} \mathcal{D} f(x,t) - r(x,t) f(x,t) = 0 \\ \text{with boundary condition } \; f(x,T) = 1 \end{gathered} \tag{2.3}$$

where $(x, t) \in \mathbb{R} \times [0, T)$, and $\mathcal{D}$ is an operator which takes the drift component of

the derivative[1]. The Feynman-Kac probabilistic solution to eq(2.3) is

$$f(x,t) = E_t\left[\exp\left(-\int_t^T r(X_s,s)ds\right)\right] \tag{2.4}$$

Now note $B(t,T,X_t) = f(x,t)$. Thus given dynamics for the stochastic process X_t, eq(2.3) provides a method to compute bond prices. Note the PDE of eq(2.3) has been restricted to reflect the specific form of eq(2.2), including the unitary boundary conditions reflecting the one dollar notational value of $B(t,T)$. For clarity we note the function dependence of B on X_t is often omitted.

To expand on eq(2.3) we require additional structure on X_t dynamics. Assume X_t is a continuous time drift-diffusion process in $\mathbb{R}^N$ with the following specification

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dW_t \tag{2.5}$$

Where dW_t is a N dimensional Brownian motion. Substituting the dynamics of eq(2.5) into eq(2.3) yields

$$f_t(x,t) + f_x(x,t)\mu(x,t) + \frac{1}{2}tr[\sigma(x,t)\sigma(x,t)^\top f_{xx}(x,t)] - r(x,t)f(x,t) = 0 \tag{2.6}$$

The solution is also expandable to include the presence of measurable jump processes. Assume X_t is a Markov process with drift, diffusion, and Poisson type jump components.

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dW_t + J(X_t,t)dN_t \tag{2.7}$$

Where N_t is a vector of Poisson jump processes, $J(X_t,t)$ are jump amplitudes, with each jump process having an associated jump intensity λ_i. In general the intensity, as well as jump amplitude may depend on X_t. Substituting the dynamics in eq(2.7)

[1] $\mathcal{D}$ is also known as the infinitesimal generator, infinitesimal operator, Dynkin operator, the Itô operator, or the Kolmogorov backward operator. See [48] for a mathematical presentation, or [4] for a more finanical context.

into eq(2.3) yields

$$f_t(x,t) + f_x(x,t)\mu(x,t) + \frac{1}{2}tr[\sigma(x,t)\sigma(x,t)^\top f_{xx}(x,t)] - r(x,t)f(x,t)$$
$$+ \lambda_i \sum_i E_i[f(x + J(x,t),t) - f(x,t)] = 0 \tag{2.8}$$

where there are i jump components. Details on admissible Poisson jump specifications, as well as sufficient conditions for existence of eq(2.8) is covered in [15].

If X_t is scalar and observable, then the PDEs of eq(2.6) and eq(2.8) may be integrated to construct bond prices. State of the art research, including the currently reported work, focuses on latent $X_t \in \mathbb{R}^N$ where $N \geq 3$. As such direct PDE integration of eq(2.6) and eq(2.8) is computationally infeasible. A well reported means to achieve a tractable solution, is to restrict X_t dynamics such that the PDEs are reduced to a system of ODEs. The reduced system of ODEs are easily solved via numerical integration. As described in section 2.2, this class of models is referred to as Affine Term Structure Models.

2.2 Affine Term Structure Models

Section 2.1 outlines the general method for constructing a pricing function for bonds. When specifying state variables, one is often faced with a trade-off between richness of dynamics and computational tractability. A well reported class of tractable models are Affine Term Structure Models (ATSM). A dynamic term structure model is affine if yields are affine in the state variables, or alternatively bond prices are exponentially affine. Tractability in ATSM models is derived from the ability to reduce the Feynman-Kac PDE into a system of ODEs, which are easily solved via the method of undetermined coefficients. As reported in [29] any state space which possess an exponentially affine conditional characteristic function may lead to an ATSM model[2].

Pioneering work in dynamic term structure modeling can be viewed as one dimensional ATSM models[59, 17]. The lack of fit to historical dynamics and conditional

[2]We also require that $r(X_t)$ be affine in X_t.

moments has encouraged the development of multi-factor or multidimensional models. Multi-factor ATSMs were originally formalized in [27, 28]. Coverage of ATSM models for X_t in the multi-factor affine drift-diffusion family is found in [19]. [29] contains a comprehensive presentation on pricing a wide range of securities which are driven by an affine jump-diffusion state space. A rigorous mathematical presentation of general affine processes with applications in finance can be found in [26]. An excellent survey paper on ATSM models is found in [52], while the survey in [18] includes expanded coverage beyond pricing risk free bonds.

Dynamics Under the Pricing Measure

We begin a brief summary of bond pricing for ATSM models with X_t defined as a drift-diffusion process. We note all processes in this section are under the Q pricing measure as defined in eq(2.3). Using the notation of [19] define X_t as

$$dX_t = \mathcal{K}(\Theta - X_t)dt + \Sigma\sqrt{S_t}dW_t^Q \tag{2.9}$$

where W_t is a N dimensional vector of independent Brownian motions, $\mathcal{K}$ and Σ are $N \times N$ matrices, Θ is a $N \times 1$ vector, and S_t is a diagonal matrix with the ith diagonal element given by

$$[S_t]_{ii} = \alpha_i + \beta_i^\top X_t \tag{2.10}$$

Assume an affine form for the short rate $r(X_t, t)$ as

$$r(X_t, t) = \delta_o + \delta_x^\top X_t \tag{2.11}$$

and a solution for Bond prices of

$$B(t, n) = exp(C_o(n) - C_x(n)^\top X_t) \tag{2.12}$$

where $n = T - t$. Substituting eq(2.9-2.12) into eq(2.6) yields ordinary differential equations (ODEs) for $C_o(n)$ and $C_x(n)$.

$$\begin{aligned}\frac{dC_o(n)}{dn} &= -\Theta^\top \mathcal{K}^\top C_x(n) + \frac{1}{2}\sum_i^N [\Sigma^\top C_x(n)]_i^2 \alpha_i - \delta_o \\ \frac{dC_x(n)}{dn} &= -\mathcal{K}^\top C_x(n) + \frac{1}{2}\sum_i^N [\Sigma^\top C_x(n)]_i^2 \beta_i - \delta_x \end{aligned} \tag{2.13}$$

with initial conditions $C_o(0) = 0$ and $C_x(0) = [0]$. For some specifications the ODEs have closed from solutions, others can be solved via numerical integration.

As discussed in [19], two key issues in specification of the any state space is admissibility and econometric identification. Admissibility is primarily concerned with ensuring that any component of X_t which has an associated nonzero β_i is nonnegative with probability 1 and thus real valued. This forces constraints within components of dX_t, as well as correlation between components. Econometric identification is concerned with ensuring unique model prices, and identification of all model parameters. A primary contribution of [19] is to formulate a canonical representation for nested families of commonly used ATSM models. This representation allows for identification of restrictive assumptions, thus giving the most flexible model possible.

Work in ATSM models when dX_t includes jump components is found in [29, 10]. The admissibility for ATSM models with state variables being defined as affine jump-diffusions (AJD) includes conditional mean and variance of dX_t, as well as the short rate are affine in X_t. To maintain tractability which is the hallmark of ATSM models, the jump intensity λ_t and jump amplitude J_t cannot both depend on X_t. Assume ATSM conditions given for drift-diffusions hold, and extend the state space to include a jump process. Define the intensity of the jump process to be affine in X_t, $\lambda_t = \lambda_o + {\lambda_x}^\top X_t$ and specify a constant deterministic jump size ν. Then eq(2.8) reduces

to a system of ODEs similar to eq(2.13)

$$\frac{dC_o(n)}{dn} = -\Theta^\top \mathcal{K}^\top C_x(n) + \frac{1}{2}\sum_i^N [\Sigma^\top C_x(n)]_i^2 \alpha_i - \delta_o - \lambda_o^n [exp(\nu(C_x)_j) - 1]$$

$$\frac{dC_x(n)}{dn} = -\mathcal{K}^\top C_x(n) + \frac{1}{2}\sum_i^N [\Sigma^\top C_x(n)]_i^2 \beta_i - \delta_x - \frac{d\lambda_x}{dC_x} exp(\nu(C_x)_j) \quad (2.14)$$

where $(C_x)_j$ is the element of C_x which corresponds to the jump process.

Change of Measure

As stated in section 2.1 we identify the measure associated with X_t as P, that is $X_t \in \mathbb{R}^N$ defined on a probability space $(\Omega, \mathcal{F}, P)$. The Q measure is a constructed measure, such that the expectation of eq(2.2) is equal to bond prices. Since the seminal work in [38, 39] arbitrage free pricing has been built on the existence of an equivalent martingale measure Q, often referred to as the risk-neutral measure. See [4] for an excellent textbook treatment on the arbitrage pricing, and [25] for a classic though more compact presentation.

Girsanov's theorem provides the machinery to construct a martingale measure which is equivalent to P. For diffusion processes, Girsonov's theorem allows us to write

$$dW_t^Q = dW_t^P + \Lambda(X_t)dt \quad (2.15)$$

for any adapted process Λ. Define X_t as a drift-diffusion under P

$$dX_t = \mu^P(X_t)dt + \sigma(X_t)dW_t^P \quad (2.16)$$

Applying eq(2.15), we find dX_t under Q

$$dX_t = \left[\mu^P(X_t) - \sigma(X_t)\Lambda(X_t)\right] dt + \sigma(X_t)dW_t^Q \quad (2.17)$$

The change in drift under Q is the mathematical mechanism allowing investors to demand additional premium on the return of an asset. For this reason $\Lambda(X_t)$ is often

referred to as the price of risk.

To maintain ATSM tractability the moments of X_t under Q must be affine in X_t, however there is no such restriction under P. Restrictions on the dynamics of X_t under P are primarily driven by the estimation scheme used to calibrate model parameters to observed data. A popular choice in the literature is to specify the dynamics of X_t as affine under Q and P. To this end [12] provides the mathematical justification for an *extended affine* price of risk[3]. The formulation essentially allows a fully flexible drift specification under P and Q. As summarized in [57], several non-affine P specifications have been reported over the years.

Similar change of measure techniques for jump-diffusions have been reported. As is typical, we define the P measurable jump intensity as an affine function of the state vector

$$\lambda_t^P = \lambda_o + \lambda_X^T X_t \tag{2.18}$$

Then we can write the Q measurable jump intensity with respect to any adapted process Λ

$$\lambda_t^Q = \lambda_t^P(X_t)\left[\Lambda_t(X_t) - 1\right] \tag{2.19}$$

Restrictions on Λ_t ensure $\lambda_t^Q \geq 0$ and is non-explosive. An overview of jump-diffusion models is found in [56], and [29] reports change of measure requirements for affine jump-diffusions.

[3][12] show under mild restrictions that Λ_t remains non-explosive, and thus is an equivalent martingale measure.

Chapter 3

DTSM of Central Bank Policy

3.1 Motivation

The Federal Reserve Bank (Fed) maintains a publicly announced target rate for overnight loans made between depository institutions. Our motivation to include the target rate within a dynamic term structure model may be categorized into two areas: economic and mathematical. We find the unique characteristics of the target rate to be easily incorporated into a dynamic term structure model.

The economic motivation to include the target rate stems from it's use as a key tool of monetary policy. The Fed provides the following description of monetary policy and the tools at it's disposal. [1].

> The term "monetary policy" refers to the actions undertaken by a central bank, such as the Federal Reserve, to influence the availability and cost of money and credit to help promote national economic goals. The Federal Reserve Act of 1913 gave the Federal Reserve responsibility for setting monetary policy.
>
> The Federal Reserve controls the three tools of monetary policy–open market operations, the discount rate, and reserve requirements. The Board of Governors of the Federal Reserve System is responsible for the

[1]Taken from the website of the Federal Reserve Bank: www.federalreserve.gov

> discount rate and reserve requirements, and the Federal Open Market Committee is responsible for open market operations. Using the three tools, the Federal Reserve influences the demand for, and supply of, balances that depository institutions hold at Federal Reserve Banks and in this way alters the federal funds rate. The federal funds rate is the interest rate at which depository institutions lend balances at the Federal Reserve to other depository institutions overnight.
>
> Changes in the federal funds rate trigger a chain of events that affect other short-term interest rates, foreign exchange rates, long-term interest rates, the amount of money and credit, and, ultimately, a range of economic variables, including employment, output, and prices of goods and services.

The Fed target rate is a publicly announced goal for the Fed funds rate. Federal Open Market Operations (FOMC) attempts to steer the daily effective funds rate toward the target rate, by supplying or withdrawing liquidity [31]. This is carried out in part by the trading desk of the Federal reserve bank in New York. Market influences and institutional constraints force temporary deviations, though on average the Fed has been very successful in keeping the rate at the intended target [9]. Figure 3.1 shows the Fed's monthly tracking error for the period of this study. Using non-overlapping months we find the mean absolute tracking error to be 2.61 basis points. The main motivation to use the target rate over the funds rate is the high volatility of the funds rate, which it shares with most short dated assets.

The combined effect of it's use as a monetary policy tool and short dated reference, results in the target acting as an anchor for longer dated yields. Figure 3.4.2 shows a time series plots for the target and a subset of synthetic yields used in our study. Except for rare times of extreme displacement, the target is seen as an anchor for longer maturity yields. Finally, several empirical studies have documented that yields of all maturities respond to unanticipated changes in the target rate [45, 53, 33]. In summary the economic motivation to include the target rate is it's use in monetary policy and the interconnected characteristic as an anchor for longer maturity yields.

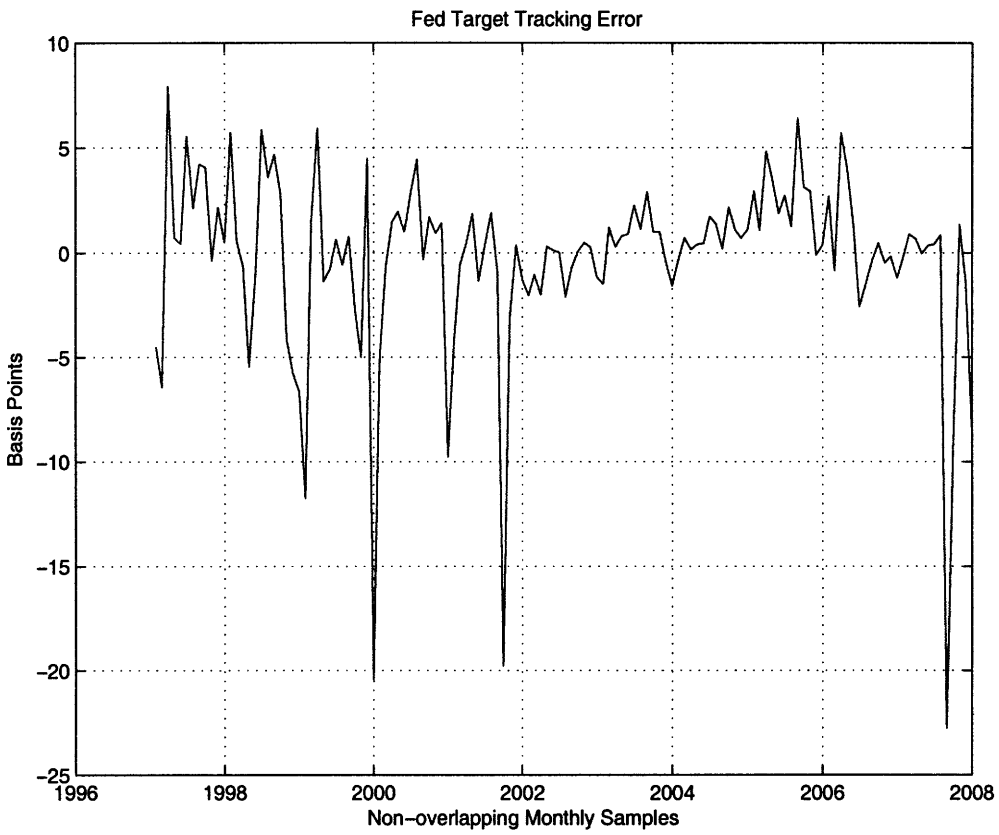

Figure 3-1: Monthly estimates of tracking error with respect to the Fed target rate. Tracking error is difference in non-overlapping monthly averages of the Fed target and the Fed effective funds rates. Sample mean absolute tracking error 2.61 bp, sample vol 4.56 bp.

From a mathematical modeling perspective, the target possesses several unique characteristics with respect to other interest rates. In 1994 the FOMC made significant changes in their operating policy in an effort to increase operational transparency. These changes include maintaining the target in 25 bp increments, as well as announcing changes during scheduled meetings. See [51] for a discussion of operational policy before 1994. Since our data is restricted to a post-1997 time period, we will focus on the *new* policy operations. Figure 3.1 shows the time series of the Fed target rate. Since 1997, 44 of the 50 target changes have occurred during scheduled FOMC meetings. Table 3.1 contains target changes which were announced during unscheduled FOMC meetings, all of which occurred during stressful economic conditions. The 1998 change is associated with the Russian financial crisis, 2001 changes with the

9/11 terrorist attack and subsequent recession, and the 2008 changes with the recent sub-prime crisis and ensuing recitation.

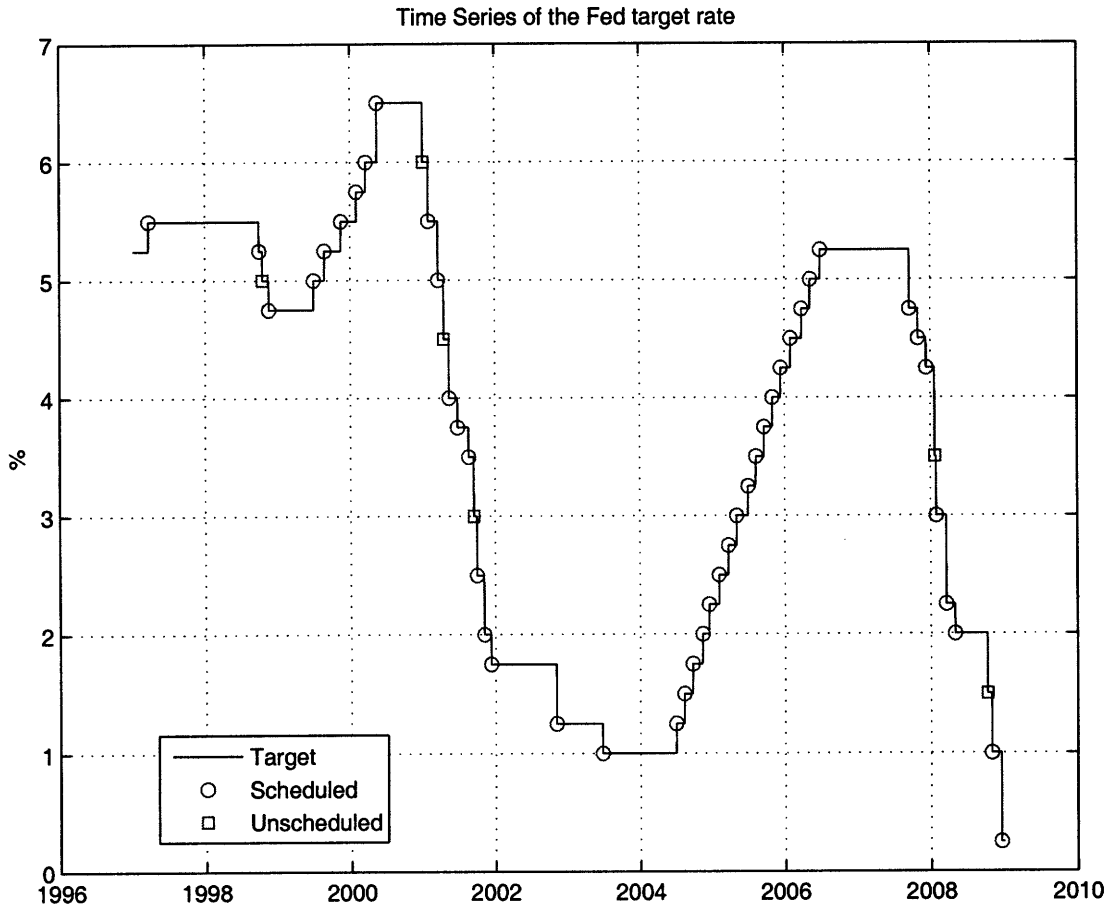

Figure 3-2: Times Series of the Fed target rate, the intended overnight Fed funds rate set by the FOMC. Changes to the target announced during scheduled FOMC meetings (circle), changes to the target announced during unscheduled FOMC meetings (square).

As seen in figure 3.1 any model constructed dynamics for the target will require discontinuous dynamics. As detailed in section 3.2.2 we select a conditionally Poisson counting process to describe the dynamics of the target. The associated jump intensity is defined with distinct dynamics for scheduled FOMC meetings and the rare jumps outside of scheduled FOMC meetings.

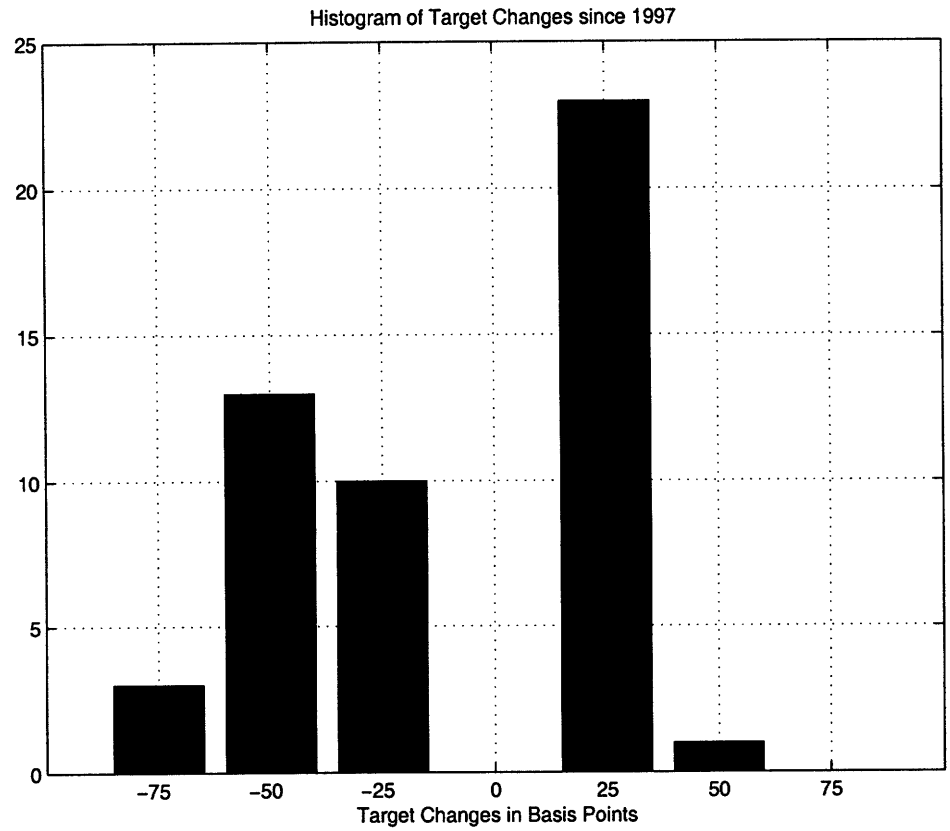

Figure 3-3: Histogram of changes to the Fed target.

Date	Changes to the Target (bp)
15 Oct 1998	-25
03 Jan 2001	-50
18 Apr 2001	-50
17 Sep 2001	-50
22 Jan 2008	-75
08 Oct 2008	-50

Table 3.1: Table of changes to the target which were announced during unscheduled FOMC meetings. Includes dates from January 1 1997 to January 1 2009.

3.2 Model Construction

In this section we describe the construction of a four factor dynamic term structure model with explicit modeling of central bank policy via the Fed target rate. The model construction closely follows that of [51]. The model possesses three latent state variables, as well as the observable target rate. The latent state variables are

continuous drift-diffusions, with stochastic volatility via a single CIR process. The observable target rate possesses a stochastic state dependent jump intensity during scheduled FOMC meetings, and a low constant intensity outside of scheduled meetings. We infer the latent states by identifying three observables yields to be error free, and identify optimal model parameters via a simulated maximum likelihood scheme with importance sampling. Finally we present the workings of a Bayesian Particle filter, which is used to verify robustness of the estimation method.

3.2.1 Latent State Space

In the seminal paper of [46], a principle component analysis of yield data reveals that three components explain the vast majority of variation in yields. Table 3.2 shows the amount of total variation explained by the first five principle components. This empirical fact has lead the research community to focus on three factor models when reporting on DTSM models [2]. This is especially true when working with latent state spaces, as additional variables present data fitting issues.

k	Yields in Levels	Changes in Yields
1	93.396	91.149
2	99.700	97.671
3	99.954	99.365
4	99.995	99.706
5	99.999	99.849

Table 3.2: Percent of total variation explained by the first k principle components, where the principle component decomposition is performed on yields (levels) and first differences of yields (changes).

Motivated by the principle component findings we construct our latent state space with three state variables, denoted by $X_t = [X_t^1, X_t^2, X_t^3]$. Each state variable is a continuous drift-diffusion process with the following dynamics

$$dX_t = \mu^P(X_t)dt + \sigma(X_t)dW_t^P \tag{3.1}$$

[2]Attempts to model specific characteristics of yields often lead to additional state variables, such as the goal of fitting very short dated yields or money market rates.

where

$$\mu^P(X_t) = K_o^P + K_1^P \cdot X_t = \begin{bmatrix} k_{01}^P \\ k_{02}^P \\ k_{03}^P \end{bmatrix} + \begin{bmatrix} k_{11}^P & 0 & 0 \\ k_{21}^P & k_{22}^P & k_{23}^P \\ k_{31}^P & k_{32}^P & k_{33}^P \end{bmatrix} \cdot X_t \tag{3.2}$$

and

$$\sigma(X_t) = \begin{bmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1 + b_{21} X_t^1} & 0 \\ 0 & 0 & \sqrt{1 + b_{31} X_t^1} \end{bmatrix} \tag{3.3}$$

where W_t^P are Wiener processes under the data generating or historical measure P. X_t^1 is a square root or CIR[3] process, which results in X_t^1 possessing time varying conditional volatility[4]. The construction of $\sigma(X_t)$ then couples the stochastic volatility to the other state variables. The off diagonal terms in the K_1^P drift component allow for full flexibility with respect to correlation between state variables. Admissibility constraints are required to ensure $X_t^1 \geq 0$, and thus $\sigma(X_t)$ remains real valued. These constraints include

1. The two zeros in the first row of K_1^P
2. $k_{01}^P \geq 0.5$
3. $b_{j1} \geq 0$ for $j = 1, 2$

In the language of [19], X_t is a $A_1(3)$ model. The notation implies only one state variable is allowed to drive instantaneous conditional volatility, where there are three state variables in total. Staying within the drift-diffusion framework and using this notation, possible model choices for X_t include $A_o(3)$, $A_1(3)$, $A_2(3)$, and $A_3(3)$. $A_o(3)$ is unique in that σ is a matrix of constants, resulting in X_t possessing a convenient joint Gaussian distribution. The positives for an $A_o(3)$ model include it's superior

[3]The seminal paper of [17] presented the dynamics for the first time in a term structure framework.

[4]Unless otherwise stated time varying volatility and stochastic volatility are used interchangeably, as is conditional versus unconditional volatility

ability to fit the yield curve and capture risk premium[5]. The overriding negative feature is the resulting constant conditional variance in model yields, which is strongly rejected by empirical studies of fixed income data. All $A_j(3)$ models possess stochastic volatility in X_t, which is then inherited by model yields. The downside of constructing stochastic volatility is admissibility constraints in the drift of the stochastic driver, i.e. the zeros in the first row of K_1^P. Such constraints decrease drift coupling, which is viewed as an important characteristic of high preforming models. Not surprisingly, the number of restrictions increases with j. The $A_1(3)$ model is chosen as the most flexible construction, which accurately reflects time varying volatility in observed yields. See [19] for one of the initial discussions on this topic, and [20] for a broader discussion on the $A_j(n)$ framework.

3.2.2 Jump Process

Following section 3.2, model dynamics for the target rate are discrete valued and move predominately during scheduled FOMC meetings. In line with [51], we select a Poisson jump process as the kernel to construct target dynamics. During scheduled FOMC meetings Poisson jumps possess a stochastic intensity driven by all state variables, while outside of scheduled meeting days jumps are driven by a small constant intensity. For convenience of presentation define the target rate as θ_t, and a superset of state variables as $\tilde{X}_t = [\theta_t, X_t]^\top$ where $X_t = [X_t^1, X_t^2, X_t^3]^\top$ as defined in section 3.2.1.

Heuristically we view X_t as key indicators of the general economy, as such they should influence the decision making process of the FOMC committee. Mathematically we implement this relationship by defining the jump intensity as an affine function of X_t. We can also view the current target rate as an additional proxy for the state of the economy. For example a historically low target rate would increase the probability that the Fed is currently attempting to expand the monetary base in order to provide credit and spur growth. This type of economic information imbedded in the target level, may or may not be contained in X_t as such we expand the jump

[5]Risk premium, bond returns, and excess return are used interchangeably. Excess return is the return one gains from holding a bond, over the promised yield available in the market.

intensity to include θ_t.

As shown in figure 3.1 the target moves in increments of 25 basis points (bp). A continuous time setting implicitly allows the jump process to register several jumps over any finite period of time, accommodating any net change equal to a multiple of 25 bp. The strict definition of a (compound) Poisson process must be extended to allow the target to increase or decrease. In [51] this is accomplished by constructing two competing Poisson process, one with positive jumps and the other with negative jumps. Unfortunately with this construction ensuring non-negative jump intensities in an affine framework is not possible. To circumnavigate this difficulty we implement a non-linear switching mechanism. We formalize this description as

$$d\theta_t = \text{sign}\left(\lambda_t\right) J_\theta dN_t^P \quad \text{for t} \in \text{scheduled FOMC meeting} \tag{3.4}$$

where the intensity of dN_t^P is equal to $\left|\lambda_t^P\right| = \left|\lambda_o^P + \lambda_\theta^P \theta_t + \lambda_X^P X_t\right|$, and J_θ is equal to 25 bp. Note when λ_t^P is positive θ_t may only jump up, and when λ_t^P is negative θ_t may only jump down. Unlike in a competing Poisson process framework, our jump intensity is strictly positive by construction. How we handle the non-linearities of eq(3.4) when constructing bond prices is discussed in section 3.2.4.

Outside of scheduled FOMC meetings we could use a similar construction as in eq(3.4). However since jumps outside of scheduled meetings are so rare, the dynamics would have to be different. On possibility is to follow eq(3.4) with λ_t^P scaled drastically downward. This would allow the state of the economy, $\tilde{X}_t$, to influence the unscheduled jumps while ensuring they are probabilistically rare. However the unscheduled jumps are so rare, as to force the scaling factor to zero. We instead choose a more parsimonious framework, allowing jumps during unscheduled meetings to occur according to a small constant intensity. Specifically outside of scheduled FOMC meetings we construct the following dynamics for θ_t.

$$d\theta_t = J_\theta \left(dN_t^u - dN_t^d\right) \quad \text{for t} \notin \text{scheduled FOMC meeting} \tag{3.5}$$

where the jump intensity of dN_t^u and dN_t^d is equal to $\overline{\lambda}$, and J_θ equals 25 basis points.

Note we do not observe dN_t^u or dN_t^d separately, rather we observe the difference.

3.2.3 Change of Measure

Recall the bond pricing relation of eq(2.2)

$$B(t,T,X_t) = E_t^Q\left[\exp\left(-\int_t^T r(X_s,s)ds\right)\right] \tag{3.6}$$

which gives model bond prices as the expectation of a functional under the Q measure. The dynamics of the state variables specified in section 3.2.1 and 3.2.2 are under the data generating or historical measure P measure. To transform eq(3.6) into a useable form we require state space dynamics under the Q measure. We address change of measure issues for the continuous latent state space X_t, and the discontinuous observable θ_t separately.

With respect to the latent state space X_t, we turn to the *extended affine* market price of risk as reported in [12]. Recall Girsanov's theorem applied to a drift-diffusion transforms the drift, but keeps the diffusion component unchanged.

$$\begin{aligned} dX_t &= \mu^P(X_t)dt + \sigma(X_t)dW_t^P & (3.7)\\ &= \mu^Q(X_t)dt + \sigma(X_t)dW_t^Q & (3.8)\end{aligned}$$

where in general

$$\mu^Q(X_t) = \mu^P(X_t) + \sigma(X_t)\Lambda(X_t) \tag{3.9}$$

where $\Lambda(X_t)$ is often referred to as the price of risk. To appreciate the significance of this label, we note for ATSM models the drift component typically dominates the pricing function. Combine this characteristic with the heuristic view of $\sigma(X_t)$ a measure of risk in state space or economy which it represents. Thus the change of measure adjusts prices by injecting a scaled measure of risk into the drift of X_t. Hence $\Lambda(X_t)$ is the price of risk. Finally we remind readers that the existence, or

requirement, of the Q measure is given by the fundamental theorem of arbitrage free pricing [38, 39].

A significant component of recent literature has focused on exploring admissible parametric forms for $\Lambda(X_t)$. Admissibility essentially focuses on ensuring the Q measure is a martingale and is equivalent to P. Recently [12] reported on a specification of $\Lambda(X_t)$, which allows the most flexible drift under P and Q. For the dynamics given in section 3.2.1, the parametric form of $\Lambda(X_t)$ is

$$\Lambda_t = \begin{bmatrix} \frac{k_{01}^P - k_{01}^Q}{\sqrt{X_t^1}} + \left(k_{11}^P - k_{11}^Q\right)\sqrt{X_t^1} \\ \frac{k_{02}^P + \left(k_{21}^Q - k_{21}^P\right) + \left(k_{22}^Q - k_{22}^P\right) + \left(k_{23}^Q - k_{23}^P\right)}{\sqrt{1+b_{21}X_t^1}} \\ \frac{k_{03}^P + \left(k_{31}^Q - k_{31}^P\right) + \left(k_{32}^Q - k_{32}^P\right) + \left(k_{33}^Q - k_{33}^P\right)}{\sqrt{1+b_{31}X_t^1}} \end{bmatrix} \tag{3.10}$$

Combining eq(3.10), eq(3.9), and eq(3.1) yields the dynamics of X_t under Q

$$dX_t = \mu^Q(X_t)dt + \sigma(X_t)dW_t^Q \tag{3.11}$$

$$\mu^Q(X_t) = K_o^Q + K_1^Q \cdot X_t = \begin{bmatrix} k_{01}^Q \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} k_{11}^Q & 0 & 0 \\ k_{21}^Q & k_{22}^Q & k_{23}^Q \\ k_{31}^Q & k_{32}^Q & k_{33}^Q \end{bmatrix} \cdot X_t \tag{3.12}$$

Admissibility of the CIR process requires the zeros in the first row of K_1^Q, as well as $k_{01}^Q \geq 0.5$. The zeros in K_o^Q are due to identification reasons. If we define the short rate as a fully flexible affine function of $\tilde{X}_t$

$$r(\tilde{X}_t) = \rho_o + \rho_{\tilde{X}_t} \tag{3.13}$$

then the zeros of K_o^Q are required to uniquely identify ρ_o. This requirement is linked to our choice of calibrating model parameters to yield data. Using alternative observables, such as derivative data, allows identification without such restrictions.

Change of measure for Poisson type jump processes is well developed, but less

covered in the financial literature. An overview of jump-diffusion models is found in [56], and [29] reports change of measure requirements for affine jump-diffusions. When focusing on a change of measure it is often convenient to transform a Poisson process to a compensated process.

$$dN_t^P = \mu_\theta^P(\tilde{X}_t)dt + dM_t^P \tag{3.14}$$

where $M_t^P = N_t^P - \int_0^t \lambda_s^P ds$, and λ_t^P is the (time varying) jump intensity of N_t^P. This decomposition allows us to write the jump process in terms of a drift term and a zero mean non-Gaussian innovation, dM_t^P. For the dynamics of section 3.2.2 the drift of the compensated process is

$$\mu_\theta^P(\tilde{X}_t) = 0 \qquad \text{for t} \notin \text{scheduled FOMC meeting} \tag{3.15}$$

$$\mu_\theta^P(\tilde{X}_t) = (\lambda_o^P + \lambda_{\tilde{X}}^P \tilde{X}_t)dt \quad \text{for t} \in \text{scheduled FOMC meeting} \tag{3.16}$$

Similar to the drift-diffusion case we construct an equivalent measure by transforming the drift of dM_t^P. Since the drift of dM_t^P is linked by construction to λ_t^P, this results in a new specification for λ_t under Q. The most flexible change of measure for the jump process results in

$$\lambda_t^Q = \lambda_o^Q + \lambda_\theta^Q \theta_t + \lambda_X^Q X_t \tag{3.17}$$

where λ_X^Q is a three dimensional row vector. There are two important conversations regarding eq(3.17). One concerning our ability to identify risk premium for the jump process, when so few jumps are observed. The other involving issues of unboundedness.

If we estimate model parameters using the full data set available, we observe 96 scheduled FOMC meetings out of 626 observations. This implies very low inference with respect to the P measurable jump intensity. Note the Q measurable jump intensity of eq(3.17) affects yields at each of the 626 observations. Furthermore during the 96 scheduled meetings, we observe only 44 target changes. This results in low inference with respect to jump risk premium. Due to the low inference we'll assume

zero risk premium for the initial models. Section 4.1 explores various risk premiums for the jump process, and explores if the data supports such specification.

3.2.4 Bond Pricing Coefficients

Within the DTSM framework developing an useable form for model bond prices is focused on solving the Feynman-Kac PDE of eq(2.8), which is associated with the conditional expectation of eq(2.2). Along with the dynamics of section 3.2.3, we require parametric forms for the short rate $r(\tilde{X}_t)$ and the bond prices $B(t,T,\tilde{X}_t)$ themselves. Given all the required ingredients we must then find a solution to the PDE of eq(2.8). If standard ATSM protocol is followed the PDE will reduce to a system of ODEs. Within our model framework the resulting ODEs must be solved via numerical integration.

Follow ATSM protocol and define the short rate as an affine function of state variables.

$$r(\widetilde{X}) = \rho\widetilde{X} = \rho_o + \rho_\theta + \rho_X X \tag{3.18}$$

where ρ_X is a three dimensional row vector. Assume an exponentially affine function for model bond prices

$$B(t,T,\widetilde{X}) = \exp\left(C_o(t,T) - C_{\widetilde{x}}(t,T)\widetilde{X}\right) \tag{3.19}$$

where $C_{\widetilde{x}}(t,T)$ is a four dimensional row vector.

We first take the case when we are not in a scheduled FOMC meeting, $t \notin$ FOMC. During this *regime* the jump process is given by (4.17), and X_t dynamics are as defined

in section 3.2.3. With these substitutions we can write eq(2.8) as

$$\begin{aligned} 0 &= \frac{\partial B(t,T,\widetilde{X})}{\partial t} + \frac{\partial B(t,T,\widetilde{X})}{\partial \widetilde{X}}\mu^{Q}(X) \\ &+ \frac{1}{2}\text{trace}\left[\frac{\partial^2 B(t,T,\widetilde{X})}{\partial \widetilde{X}^2}\sigma(X)\sigma(X)^T\right] - r(\widetilde{X}) \\ &+ \bar{\lambda}\left[B(t,T,X,\theta+J_\theta) - B(t,T,\widetilde{X})\right] \\ &+ \bar{\lambda}\left[B(t,T,X,\theta-J_\theta) - B(t,T,\widetilde{X})\right] \end{aligned} \tag{3.20}$$

Expanding all terms in (3.20) results in a PDE which can be expressed as a an affine function of $\tilde{X}$. Since eq(3.20) must hold for all values of $\tilde{X}$, each coefficient in the affine representation must separately equal zero. This is the basis of the often quoted method of undetermined coefficients. The five ODEs which result from eq(3.20) are

$$\begin{aligned} \frac{dC_o}{dt} &= \rho_o - C_{X1}K_{01} - \frac{1}{2}\left(C_{X2}^2 + C_{X3}^2\right) + \bar{\lambda}\left[2 + \exp(J_\theta C_\theta) - \exp(-J_\theta C_\theta)\right] \\ \frac{dC_\theta}{dt} &= \rho_\theta \\ \frac{dC_{X1}}{dt} &= \rho_1 + C_{X1}K_{11} + C_{X2}K_{21} + C_{X3}K_{31} - \frac{1}{2}\left(C_{X1}^2 + b_{21}C_{X2}^2 + b_{31}C_{X3}^2\right) \\ \frac{dC_{X2}}{dt} &= \rho_2 + C_{X2}K_{22} + C_{X3}K_{32} \\ \frac{dC_{X3}}{dt} &= \rho_3 + C_{X2}K_{23} + C_{X3}K_{33} \end{aligned} \tag{3.21}$$

We solve the system of ODEs numerically, via the Runge-Kuttta Method[6]. Specifically the integration is started at $t = T$, where $C_j(T,T) = 0$ for all j and continued until $t = 0$.

For time during scheduled FOMC meetings, $t \in$ FOMC, the jump dynamics

[6] $\frac{dC_\theta}{dt}$ can be easily solved analytically.

change to reflect the now stochastic jump intensity of eq(3.4). The resulting PDE is

$$\begin{aligned} 0 &= \frac{\partial B(t,T,\widetilde{X})}{\partial t} + \frac{\partial B(t,T,\widetilde{X})}{\partial \widetilde{X}}\mu^{Q}(X) \\ &+ \frac{1}{2}\text{trace}\left[\frac{\partial^2 B(t,T,\widetilde{X})}{\partial \widetilde{X}^2}\sigma(X)\sigma(X)^T\right] - r(\widetilde{X}) \\ &+ \Big|\lambda(X)\Big|\Big[B(t,T,X,\theta+\text{sign}(\lambda(X))J_\theta) - B(t,T,\widetilde{X})\Big] \end{aligned} \tag{3.22}$$

Unfortunately the non-linear terms originating from the jump term prevent expressing eq(3.22) as an affine function of $\tilde{X}$. To maintain the tractability which is the hallmark of ATSM models, we linearize the jump term of eq(3.22). Specifically we apply a Taylor Series expansion

$$\begin{aligned} \Big|\lambda(X)\Big|\Big[B(t,T,X,\theta+\text{sign}(\lambda(X))J_\theta) - B(t,T,\widetilde{X})\Big] \approx \\ J_\theta\lambda(\widetilde{X})C_\theta(t,T)B(t,T,\widetilde{X}) \end{aligned} \tag{3.23}$$

which is affine in $\widetilde{X}$ since $\lambda(\widetilde{X})$ is affine by construction. With this approximation we are able to reduce eq(3.22) to the following system of ODEs for $t \in$ FOMC

$$\begin{aligned} \frac{dC_o}{dt} &= \rho_o - C_{X1}K_{01} - \frac{1}{2}\left(C_{X2}^2 + C_{X3}^2\right) - J_\theta\lambda_o C_\theta \\ \frac{dC_\theta}{dt} &= \rho_\theta - J_\theta\lambda_\theta C_\theta \\ \frac{dC_{X1}}{dt} &= \rho_1 + C_{X1}K_{11} + C_{X2}K_{21} + C_{X3}K_{31} - \frac{1}{2}\left(C_{X1}^2 + b_{21}C_{X2}^2 + b_{31}C_{X3}^2\right) - J_\theta\lambda_1 C_\theta \\ \frac{dC_{X2}}{dt} &= \rho_2 + C_{X2}K_{22} + C_{X3}K_{32} - J_\theta\lambda_2 C_\theta \\ \frac{dC_{X3}}{dt} &= \rho_3 + C_{X2}K_{23} + C_{X3}K_{33} - J_\theta\lambda_3 C_\theta \end{aligned} \tag{3.24}$$

The effect of the linearization in eq(3.23) is quantitatively measured in Appendix A. Unless otherwise noted the approximation is seen to have no noticeable effect on results, or conclusions drawn from results.

To construct pricing formulas we use the public meeting schedule of the FOMC to alternate between eq(3.21) and eq(3.24). Begin with the boundary condition

$\{C_o(T,T) = 0, C_{\tilde{X}}(T,T) = 0\}$, and integrate backwards in time until a FOMC meeting is scheduled. Use the integration result of eq(3.21) as the initial condition to the integration of eq(3.24). Continue alternating between systems of ODEs until $t = 0$, at which point $\{C_o(t,T), C_{\tilde{X}}(t,T)\}$ is available.

Unfortunately the FOMC meeting schedule changes each year, and is not uniform within the calender year. This irregularity of the meeting schedule forces us to resolve the ODEs at each observation and for each bond maturity. Within the context of an estimation scheme, the resulting computational burden is essentially infeasible. To alleviate the burden, we follow [51] and assume a uniform meeting schedule after the first meeting for each observation. The FOMC meeting schedule is known one to two years in advance, thus for market participates are not aware of the schedule for longer maturities. This fact lends support to normalize the meeting schedule. Appendix A quantifies the effect of this approximation, which is minimal for the maturities in our data set.

3.3 Estimation

A maximum likelihood scheme is chosen to estimate model parameters. Since three state variables are latent, we identify three observables as error free allowing us to infer the value of the latent variables by inverting the measurement equation. Since the transitional density of the state vector is not available in closed form, a simulated likelihood technique is employed. To speed up the convergence of the Monte Carlo integration an importance sampling density is exploited. Finally, to verify the estimation results are robust to the inferred latent state variables, a Bayesian Particle Filter is applied using the optimal model parameters.

We identify weekly sampled six month libor, and $\{1, 2, 3, ..., 9, 10\}$ year swap contracts as observables. To construct bond prices from the non-linear swap contracts we keep unobserved forward rates constant, resulting in synthetic yields, for $\{.5, 1, 2, ..., 9, 10\}$ year maturities. The benefit of synthetic yields is a measurement equation linear in $\tilde{X}$.

The log-likelihood function we seek to maximize is

$$\sum_{t=1}^{T} f_Y(y_t|y_{t-1};\gamma) = \sum_{t=1}^{T} f_{\tilde{X}}(g(y_t,\gamma) \mid g(y_{t-1},\gamma);\gamma)\Big|\nabla g(Y_t,\gamma)\Big| \tag{3.25}$$

where γ is a vector of unknown model parameters, $\tilde{X}_t = g(Y_t,\gamma)$ is the inverted measurement equation, $f_{\tilde{X}}(\cdot|\cdot)$ is the log conditional density of $\tilde{X}$, and $|\bigtriangledown g(Y_t,\gamma)|$ is the Jacobian of the measurement equation. Unfortunately the dynamics of $\tilde{X}_t$ do not permit a closed from expression[7] for $f_{\tilde{X}}(\cdot|\cdot)$. To overcome this obstacle, simulated maximum likelihood is employed.

3.3.1 Simulated Maximum Likelihood with Importance Sampling

Simulated maximum likelihood (SML) is a popular means to estimate model parameters of a continuous time stochastic process using discretely sampled data. The general idea of SML is to approximate the true conditional density by discretizing the SDEs of eq(3.1,4.17,3.4) using a Euler approximation and employing Monte Carlo to estimate the value of the conditional density. Details of the SML technique as presented by [49] can be found in Appendix B.1. The estimation scheme implemented is of the SML flavor, with variance reduction achieved via an importance sampling technique.

View the conditional density $f(y_t|y_{t-1};\gamma)$ as a marginal density, with a corresponding joint density which explicitly depends on the value of y_s for $s \in (t-1,t)$. That is

$$f(y_t|y_{t-1}) = \int f(y_t, y_t^*|y_{t-1})dy_t^* \tag{3.26}$$

where the γ dependency has been dropped for notational simplicity. If a change of measure is implemented via an importance sampling density $f_S(y_t^*)$, eq(3.26) can be

[7]Closed form solutions are known for a few special cases, namely Gaussian and square root diffusions.

written as

$$f(y_t|y_{t-1}) = \int \frac{f(y_t, y_t^*|y_{t-1})}{f_s(y_t^*)} f_s(y_t^*) dy_t^* \tag{3.27}$$

Evaluate eq(3.27) using Monte Carlo to find,

$$\hat{f}^M(y_t|y_{t-1}) = \frac{1}{J}\sum_{j=1}^{J} \frac{f(y_t, y_{t,j}^*|y_{t-1})}{f_s(y_t^*)} \tag{3.28}$$

As noted in [32], SML can now be viewed as setting the sampling density to $f(y_t^*|y_{t-1})$, thus

$$\hat{f}^M(y_t|y_{t-1}) = \frac{1}{J}\sum_{j=1}^{J} \frac{f(y_t, y_{t,j}^*|y_{t-1})}{f(y_t^*|y_{t-1})} = \frac{1}{J}\sum_{j=1}^{J} f(y_t|y_t^*, y_{t-1}) \tag{3.29}$$

Though conceptually and numerically simple this sampling density does not exploit the known realization of y_t. A more efficient sampling density is based on the work introduced in [35]. Let $\hat{\mu}_t$ be the mode of $logf(y_t, y_t^*|y_{t-1})$, and $\hat{\Sigma}_t$ be the negative of the Hessian of $logf(y_t, y_t^*|y_{t-1})$ evaluated at it's mode. The proposed importance sampling density is a multivariate Student-T density with mean μ_t and dispersion Σ_t. We denote this sampling density as $f_T(y_t^*|\hat{\mu}_t, \hat{\Sigma}_t, \nu)$, where ν indicates the degrees of freedom. The importance sampling density estimator is then

$$\hat{f}^M(y_t|y_{t-1}) = \frac{1}{J}\sum_{j=1}^{J} \frac{f(y_t, y_{t,j}^*|y_{t-1})}{f_T(y_{t,j}^*|\hat{\mu}_t, \hat{\Sigma}_t, \nu)} \tag{3.30}$$

$$y_{t,j}^* \sim f_T(y_t^*|\hat{\mu}_t, \hat{\Sigma}_t, \nu) \tag{3.31}$$

which completes the importance sampling scheme. The use of a Student-T density provides fat tails, with ν giving the means to ensure the sampling density has adequate coverage in the tails as compared to the true density. If the true density $f(y_t|y_{t-1})$ does not have significant mass in the tails, a multivariate normal may be used in place of the Student-T. A comparative empirical study on SML techniques using importance sampling is found in [30] [8]. For a drift-diffusion state space, as specified in section 2.2, the only remaining issue is the discritization used between observations which is the

[8] [30] is accompanied by an informative correspondence amongst the principle authors, and the original authors of each technique.

integer value of M in appendix B.1. For weekly observed data, [32] report empirical support for using daily discritization, that is $M = 7$.

Our model has the added complication of one state variable possessing pure jump dynamics. One option is to follow [51] by expanding the SML of appendix B.1 to model $d\theta_t$ with Bernoulli random variables. We instead use a solution which leverages the fact that θ_t is fully observable at a daily frequency. Note the conditional density of $\tilde{X}_t$ can be conveniently split into a continuous and discontinuous component. Specifically

$$f_{\tilde{X}}(\tilde{x}_t|\tilde{x}_{t-1}) = f_X(x_t|x_{t-1}) f_\theta(\theta_t|x_{t-1}, \theta_{t-1}) \tag{3.32}$$

For observations where there is no scheduled FOMC meeting $\in (t, t-1]$ the jump term is defined by eq(4.17). Due to the lack of available data we fix $\overline{\lambda}$ to it's historical value. Fixing $\overline{\lambda}$ means that $f_\theta(\theta_t|x_{t-1}, \theta_{t-1})$ is not a function of γ, and thus does not influence eq(3.25). Thus our importance sampling scheme for $\hat{f}_{\tilde{X}}(\tilde{x}_t|\tilde{x}_{t-1})$ is

$$\hat{f}_{\tilde{X}}(\tilde{x}_t|\tilde{x}_{t-1}) = \frac{1}{J}\sum_{j=1}^{J} \frac{f(x_t, x^*_{t,j}|x_{t-1})}{f_T(x^*_{t,j}|\hat{\mu}_t, \hat{\Sigma}_t, \nu)} \tag{3.33}$$

For observations where there is a scheduled FOMC meeting $\in (t, t-1]$ the jump term is defined by eq(3.4). Since the intensity is a function of $\tilde{X}$, it will affect the maximization of eq(3.25). Fix the length of any scheduled FOMC meeting to one day, and observe θ_t at a daily frequency. Focusing on days with a scheduled FOMC meeting we find eq(3.30) becomes

$$\hat{f}_{\tilde{X}}(\tilde{x}_t|\tilde{x}_{t-1}) = \frac{1}{J}\sum_{j=1}^{J} \frac{f(x_t, x^*_{t,j}|x_{t-1})}{f_T(x^*_{t,j}|\hat{\mu}_t, \hat{\Sigma}_t, \nu)} \left[\exp\left(-|\lambda(x^*_{s,j}, \theta_s)|\right) \frac{\lambda(x^*_{s,j}, \theta_s)^k}{k!}\right] \tag{3.34}$$

where k equals the number of 25 basis point jumps during the scheduled FOMC meeting, and s is the time index directly before the day of the scheduled meeting. Note eq(3.34) is implicitly holding $\tilde{X}$ constant during the actual FOMC meeting.

3.3.2 Particle Filtering

A popular sequential Monte Carlo (SMC) estimation technique is the particle filter (PF). Particle filters were first introduced by [37] as a means to estimate a nonlinear non-Gaussian latent state space. Though particle filtering is well reported as a means to infer latent states, the use of the technique for model parameter estimation is an open topic of research. In the current reported research we employ particle filtering as a means to verify robustness in the choice of yields used to invert the measurement equation.

For a gentle textbook introduction to the concepts of sequential bayesian inference and particle filters in particular see [55]. An extensive review of recent research on Bayesian Filters is found in [11], and with a focus on particle filters in [2]. The edited volume of [22] contains text book summaries of recent research, including extensions for parameter estimation. For a dense text book treatment of nonlinear filtering and estimation of latent states see [8]. A brief review of particle filters applied to financial econometric problems can be found in the survey paper of [44].

Employing Bayes rule we can express the conditional density of X_t given all observations up to time t or $Y_{1:t}$ as

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \tag{3.35}$$

In the language of filtering: $p(x_t|y_{1:t})$ is the filter density, $p(x_t|y_{1:t-1})$ is the predictive density, and $p(y_t|x_t)$ is the likelihood density. The predictive density is linked to the filter density via the state transition density $p(x_t|x_{t-1})$ as

$$p(x_t|y_{t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{t-1})dx_{t-1} \tag{3.36}$$

The signature of any particle filter is to approximate the continuous distribution $p(x_t|y_{1:t})$ as a discrete probability mass function (PMF). Each particle is a point of mass in the PMF,

$$\hat{p}(x_t|y_{1:t}) = \sum_{i=1}^{N} \delta(x_t - x_t^i)\pi_t^i \tag{3.37}$$

where N is the number of particles, and δ is the Dirac delta function. Substituting eq(3.37) into eq(3.36) allows us to evaluate the integral as a summation.

$$\hat{p}(x_t|y_{1:t-1}) = \sum_{i=1}^{N} p(x_t|x_{t-1}^i)\pi_{t-1}^i \tag{3.38}$$

Substituting eq(3.38) into eq(3.35) yields

$$\hat{p}(x_t|y_{1:t}) \propto p(y_t|x_t) \sum_{i=1}^{N} p(x_t|x_{t-1}^i)\pi_t^i \tag{3.39}$$

where the typical $p(y_t|y_{1:t-1})$ integration constant is omitted, as it is easily estimated via normalization. The idea behind particle filters is to use eq(3.39) to draw samples of $\hat{p}(x_t|y_{1:t})$. The strength of the technique is the novel method in which these draws are taken. We will focus our attention on the Sampling Importance Resampling (SIR) algorithm of [37].

SIR Particle Filter: given particles and weights at time $t-1$: $[\{x_{t-1}^i, w_{t-1}^i\}_{i=1}^N]$

1. Draw $x_t^i \sim p(x_t|x_{t-1}^i)$
2. Calculate $w_t^i = p(y_t|x_t^i)$
3. Normalize weights; $\tilde{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i}$
4. Resample particles x_t^i based on $\tilde{w}_t^i$ to yield $[\{x_t^i, w_t^i\}_{i=1}^N]$, where $w_t^i = \frac{1}{N}$.

Repeat for $t+1, t+2, \ldots, T$

See [11] for a summary of reported resampling techniques, including systematic resampling which is used in this reported work. Note the SIR algorithm only requires us to draw from the transitional density, and to evaluate the likelihood density. Given dynamics for a state space we can (almost) always take draws by the transitional density, even though evaluation of the density may be computationally prohibitive or infeasible. The likelihood function, as defined above, is given solely by our mea-

surement equation and based largely on the assumed (additive) error. Thus particle filters are extremely flexible, and appropriate for our state space construction.

As with other filtering schemes many metrics of interest are conveniently available. Estimates of latent X_t are available from the filtered density $p(x_t|y_t)$, typically the expected value. The Monte Carlo estimate of the likelihood function is

$$\begin{aligned} \mathrm{L}(y_t|y_{1:t-1}) &= \int p(y_t|x_t)p(x_t|y_{t-1})dx_t && (3.40) \\ &\approx \sum_{i=1}^{N} p(y_t|x_t^i)w_t^i && (3.41) \\ &\approx \frac{1}{N}\sum_{i=1}^{N} p(y_t|x_t^i) && (3.42) \end{aligned}$$

Unfortunately the likelihood estimate has a high amount of variance due to the resampling step. This variance results in a likelihood which is discontinuous with respect to γ. Discontinuous likelihoods mean gradient based optimization routines cannot be used.

3.4 Market Data

Model parameter estimation, as well as model performance evaluation, requires several pieces of market data. To create market implied zero yields we require LIBOR rates and swap rates. The model uses the Fed target rate as an observable state variable, and requires the FOMC meeting schedule to construct the pricing equation. We will also have need for a metric of market implied expected target changes, which is derived from Fed future contracts. The data collected spans almost 11 years, from January 7th 1997 to December 30th 2008. The lower date being bounded by the author's access to swap rates, though institutional changes would limit the data to 1994 if available.

Ideally all observations would be recorded with a uniform time and date stamp. Unfortunately, with the author's current data sources, this is not possible. As such we implement time shifts and minimize temporal misalignment by sampling at a weekly

frequency. In summary we shift all data to the London market perspective.

3.4.1 FOMC Data

The estimation scheme of section 3.3.1 requires daily sampling of the Federal Reserve target rate, while the pricing function requires weekly samples. Since 1994 changes to the target are announced during the afternoon in New York, typically at 2:15PM EST. A time series of Fed target rates is available from DataStream[21], where changes to the target are time stamped the day they are announced. To link the target rate to the London time stamp of yields, we increment all target dates by one business day. This eliminates the error with target rates as the rate recorded at 5:00PM EST is the prevailing target rate used for all trading activity on the following day in the London market. Alternatively one may note that when changes to the target are announced at 2:15PM EST, the London Market is closed, hence this change will not be reflected in the LIBOR or swap rate of the same date, but rather the rate of the following business day.

The dates of all FOMC meetings since 1970 are available via the Federal Reserve, including a tentative schedule for the following two years[9]. Scheduled FOMC meetings last one or two days, though any change in the target rate is almost always announced in the afternoon of the second day. Unscheduled meetings occur when the the FOMC announces a target change, and there is no scheduled meeting. Hence unscheduled meetings are easily identified by comparing changes in the target time series and the meeting schedule. Consistent with incrementing the target time series by one business day, we increment scheduled FOMC meetings by one business day. See [45] and [33] for additional information regarding unscheduled target changes.

3.4.2 Libor & Swaps

Zero yields are constructed from a six month libor rate, as well as $\{1, 2, 3, ..., 9, 10\}$ USD swap rates. Libor is obtained from DataStream, who in turn receive the rate

[9]In particular the Federal Reserve Bank of Minneapolis [6] maintains a website with easily accessible data.

from the British Bankers Association (BBA). The LIBOR rate used is an average price from major Banks in the London Market, recorded at 11:00AM London local time[10]. The relationship between a zero bond $B(t,T)$ and a Libor rate $L(t,T)$ is

$$B(t,T) = \frac{1}{1 + L(t,T)\tau(t,T)} \tag{3.43}$$

where $\tau(t,T)$ is the day count over $T-t$, approximately 0.5 for a six month rate. Swap rates are obtained from DataStream, who in turn receive the rate from ICAP. The swap rates are USD dominated, recorded at 5:00PM London local time, and are priced to reset twice a year (semi-annual). The relationship between a zero bond $B(t,T)$ and a semi-annual swap rate $S(t,T)$ is

$$1 = B(t,T) + \frac{S(t,T)}{2} \sum_{j=1}^{2\tau} B(t, t + \frac{j}{2}) \tag{3.44}$$

The left hand side of eq(3.44) is the present value of the floating leg of the swap, which is always equal to the notional value of the underlying. The right hand side represents the present value of the known fixed payments over the life of the swap. To construct the (synthetic) zero yield curve we bootstrap mid-year zero bonds by holding forward rates constant. See [41] for a textbook presentation on bootstrapping techniques, including constant forward rate methods.

3.4.3 Fed Future Contracts

A 30-day Fed funds future contract (Fed future) requires delivery of the monthly average Fed funds interest rate paid on a principal amount of $5 million. In practice the contracts are marked to market, or cash settled on a daily basis avoiding any payments at the contract's expiration. Fed futures settle on the last trading day of

[10]DataStream gives the following description, "The BBA LIBOR Fixing is based upon rates supplied by BBA LIBOR Contributor Panel Banks. An individual BBA LIBOR Contributor Bank contributes the rate at which it could borrow funds, were it to do so by asking for and then accepting inter-bank offers in reasonable market size, just prior to 11:00 hrs. Contributor rates are ranked in order and the middle two quartiles averaged arithmetically. Such average rate will be the BBA LIBOR Fixing for that particular currency, maturity and fixing date".

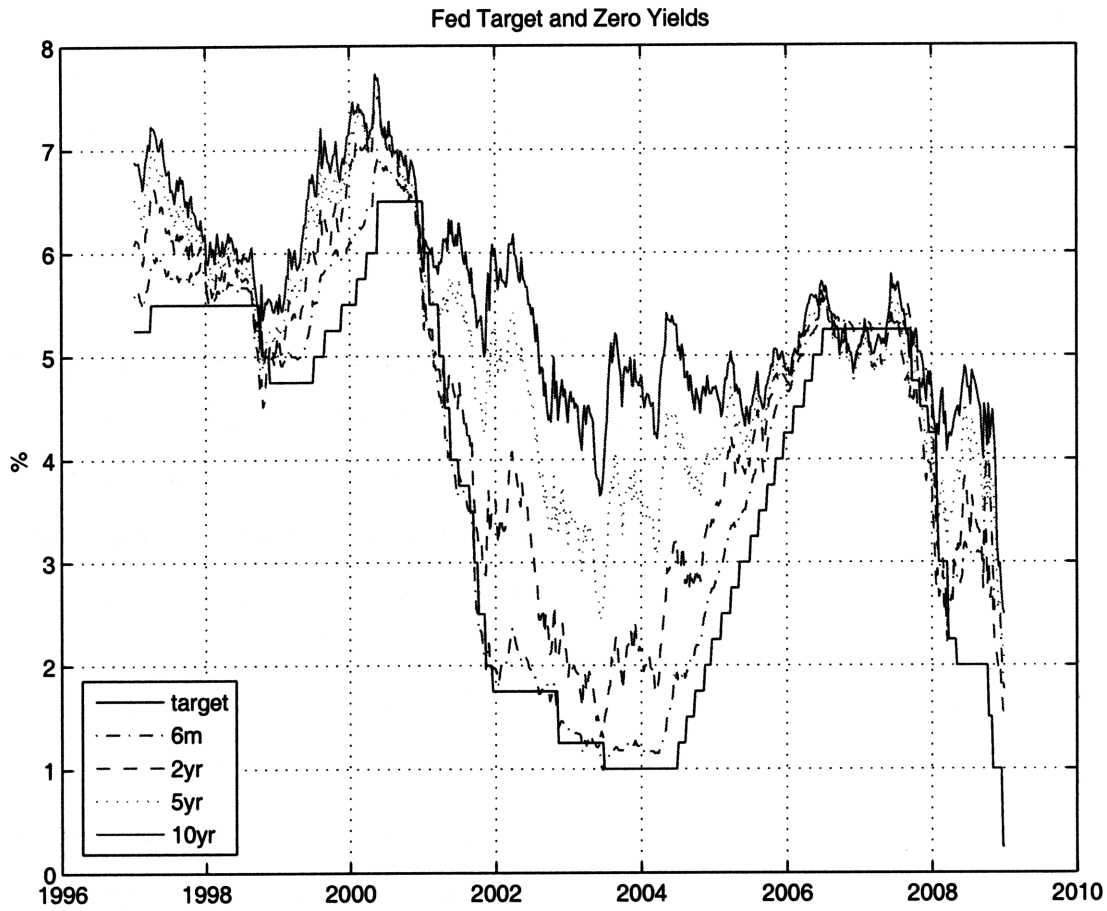

Figure 3-4: Time series of the Fed target and a subset of synthetic zero yields.

the month in which they expire, with a settlement rate[11] equal to the arithmetic average of the funds rate for the expiry month [12]. Contracts are actively traded for the current month or spot month, as well as 1 to 2 months ahead. Contracts up to 24 months in maturity do exist, though the volume is usually thin and often zero [9].

Fed fund futures are convenient tools to measure the market's expectation of changes to the target rate. An illustrative example follows. Assume you buy the spot month Fed future contract on 3/21/2002, which expires in April 2002. The target rate on 3/21/2002 was 1.75%, and the next scheduled FOMC meeting was on 5/7/2002. The futures contract should be approximately equal to $\$100 - 1.75 = \$98.25$, which indicates the market's expectation that the target rate will not change in April 2002.

[11]Contracts are priced as 100 minus the settlement rate.

[12]For weekend and holidays the rate on the previous trading day is carried over

If there is at most one FOMC meeting per month, Fed futures can be used to extract the market's expectation of target changes at FOMC meetings.

In general the time t price of the spot Fed future contract, denoted by f_t^m, is given by

$$f_t^m = 100 - \left\{ \frac{1}{t} \sum_{i=1}^{t} r_i + \frac{1}{m-t} E_t[\sum_{i=t+1}^{m} r_i] \right\} \tag{3.45}$$

where r_i is the daily Fed funds rate and m is the total number of days in the month. The Fed has historically been able to keep the average Fed funds rate within basis points of the target [36]. This documented fact leads one to approximate any finite sample mean of the Fed funds rate by the target rate over the sample. If one then assumes the target is moved only during a FOMC meeting, eq(3.45) provides a convenient means to estimate the market's expectation with respect to changes in the target rate.

A few points are worth mentioning. The right hand side of eq(3.45) could also contain a risk premium, to compensate a person for holding the risk associated with the next $m - t$ days. [53] and [13] find evidence of a small risk premium on the order of a few basis points at a one month horizon. Also note that eq(3.45) is for the current spot month. If t is in the 2nd half of the month, there is a meeting before the end of the month, and there is no meeting scheduled for the following month the one month ahead future is used instead of the spot month. There are two main reasons for this strategy. As $t \rightarrow m$ target changes will have less of an effect on the months average, and daily fluctuations in the Fed funds rate will add noise to the estimate of the expectation. It's also been noted that spot contract volume decreases drastically as $t \rightarrow m$, which reduces the information content of the contract [9], [1].

Chapter 4

Results & Performance

In this chapter we report on model results and performance. We begin with estimation results, including point estimates for all free model parameters. We present a full characterization of the associated state space, with emphasis on contributions to the short rate and jump intensity. We include a brief summary of alternative model configurations and the resulting implications. We provide in and out-of-sample pricing errors, with some surprising results. We document yield responses to innovations in the state space, comparing our results to existing literature. Finally we document the model's ability to reflect risk premium. We conclude the addition of the target greatly enhances the model's ability to describe risk premium. The inclusion of the term structure of target rates is seen to extend this performance.

4.1 Estimation Results

To construct point estimates for all model parameters we employ the simulated maximum likelihood (SML) technique as described in section 3.3.1. Specifically we construct a log likelihood function and use a gradient based optimization routine to identify a global maximum[1].

To infer values of the latent state space we assume three yields are observed without error, and invert the pricing equation. The remaining yields are constructed

[1]A flexible non-linear gradient based optimization routine in MATLAB is fmincon.m.

to possess an additive error

$$Y(t,T) = C_o(t,T) + C_{\tilde{X}}(t,T)\tilde{X}_t + \epsilon_t \tag{4.1}$$

where ϵ_t is a eight dimensional zero mean multivariate Gaussian noise term with a diagonal covariance matrix of equal entries. Following this construction a single term defines the distribution of ϵ_t, which we denote σ_ϵ. σ_ϵ is treated as a nuisance parameter and integrated out during each evaluation of the likelihood function. We also compare these estimates of X_t with ones obtained from the Bayesian Particle Filter of section 3.3.2.

Model Parameter Estimates

Table 4.1 reports SML estimates using weekly samples from January 1997 to January 2007 totalling 521 samples. We chose to remove recent data from the calibration period, as the *sub-prime crisis* introduced significant credit risk into LIBOR and Swap contracts. From January 1997 to January 2007 there were four unscheduled FOMC meetings with a total of seven 25 basis point jumps, see table 3.1 for details. This extremely rare occurrence of target changes during unscheduled FOMC meetings, encourages us to fix $\overline{\lambda}$ to it's historical level of 0.7019. To reflect the target's important role in the short rate we also fix ρ_θ to unity. Unfortunately armed with the short sample period we are unable to estimate all model parameters with statistical significance. Specifically the cross terms in the drift of $d\tilde{X}_t$, under both P and Q, are estimated without significance.

Estimated latent state variables are displayed in figures 4-1-4-3. In each case the particle filter estimates are very close to the more ad hoc method of assuming three yields are error free. Indicating a lack of tension in choosing specific yields to be observed without error. Unless otherwise stated results are identical using either inference method.

Common to almost all DTSM models there exists one highly persistent state variable under Q. Table 4.2 reports the half-life of shocks for an equivalent diag-

Parameter	Estimate	Standard Error	t-Statistic
k_{o1}^{Q}	292.92	1079.13	0.271
k_{11}^{Q}	1.31	0.06	**23.101**
k_{21}^{Q}	0.46	7.14	0.065
k_{22}^{Q}	2.24	0.09	**24.513**
k_{23}^{Q}	-0.39	6.95	-0.056
k_{31}^{Q}	-1.71*1e-3	0.03	-0.053
k_{32}^{Q}	0.04	0.66	0.056
k_{33}^{Q}	0.18*1e-3	0.01	0.013
b_{21}	0.055	1.85	0.030
b_{31}	0.063	2.41	0.026
ρ_o	-0.10	0.17	-0.566
ρ_θ	1.00	...	...
ρ_1	0.19*1e-3	0.00	0.535
ρ_2	0.01*1e-6	0.00	0.003
ρ_3	0.23*1e-3	0.00	0.052
λ_o	0.51	236.00	0.002
λ_θ	-9941.19	64.87	**-153.254**
λ_1	5.18	9.87	0.525
λ_2	141.16	2386.06	0.059
λ_3	-0.01	2.89	-0.004
k_{o1}^{P}	690.85	2558.39	0.270
k_{o2}^{P}	-0.32	203.07	-0.002
k_{o3}^{P}	437.55	8949.34	0.049
k_{11}^{P}	3.10	0.85	**3.662**
k_{21}^{P}	0.08	1.27	0.061
k_{22}^{P}	0.97	0.48	**2.026**
k_{23}^{P}	-0.05	1.09	-0.048
k_{31}^{P}	0.40	7.24	0.056
k_{32}^{P}	0.11	2.07	0.055
k_{33}^{P}	1.46	0.64	**2.293**
$\bar{\lambda}$	0.7019	...	...

Table 4.1: Point parameter estimates for model parameters of section 3.2. Estimates are calculated via the simulated maximum likelihood technique described in section 3.3.1. Sample period is from January 1997 to January 2007, including 521 weekly samples. Standard errors are computed via the product of outer gradients [3].

onalized system, which reveals our model shares this common trait. X_t^1 is not a candidate to carry the persistent shocks as it's time constant is directly calculated as $-ln(0.5)/k_{11}^{Q} = 0.5221$. An inspection of table 4.1 reveals an ultra low speed of

mean reversion for X_t^3 under Q. This low value for k_{33}^Q, indicates the highly persistent variable is mostly carried by X_t^3. This is also supported by the fact that X_t^3 has the largest effect on long maturity yields, as reported in section 4.3. Finally we note that X_t^2 has a significant shift in it's long run mean between the two measure. This trait leads to it's significant contribution to explaining risk premium as reported in section 4.4.

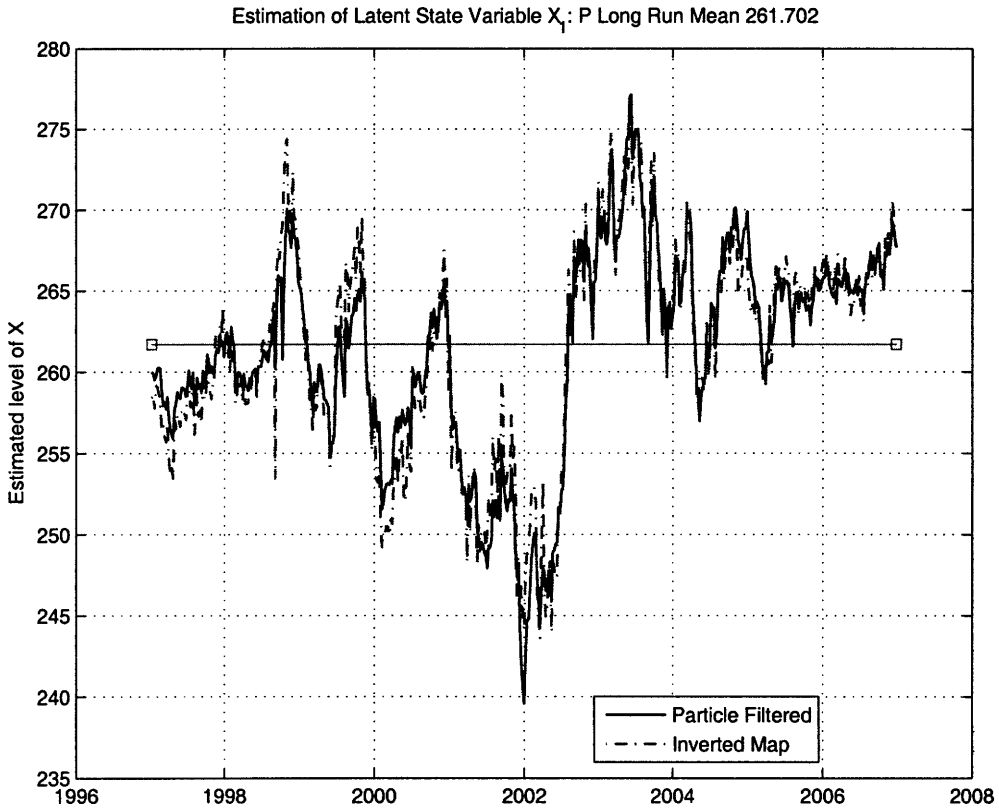

Figure 4-1: Estimated values of the stochastic volatility factor X_t^1. Horizontal bar fixed at the P measurable long run mean of X_t^1.

Q Half-Life	P Half-Life
93.6198	.4657
0.3115	0.7142
0.5221	0.2252

Table 4.2: Half-life of shocks for an equivalent diagonalized system of X_t. Specifically $-log(0.5)\ \Lambda^{-1}$, where Λ are the eigenvalues associated with the drift matrix of X_t.

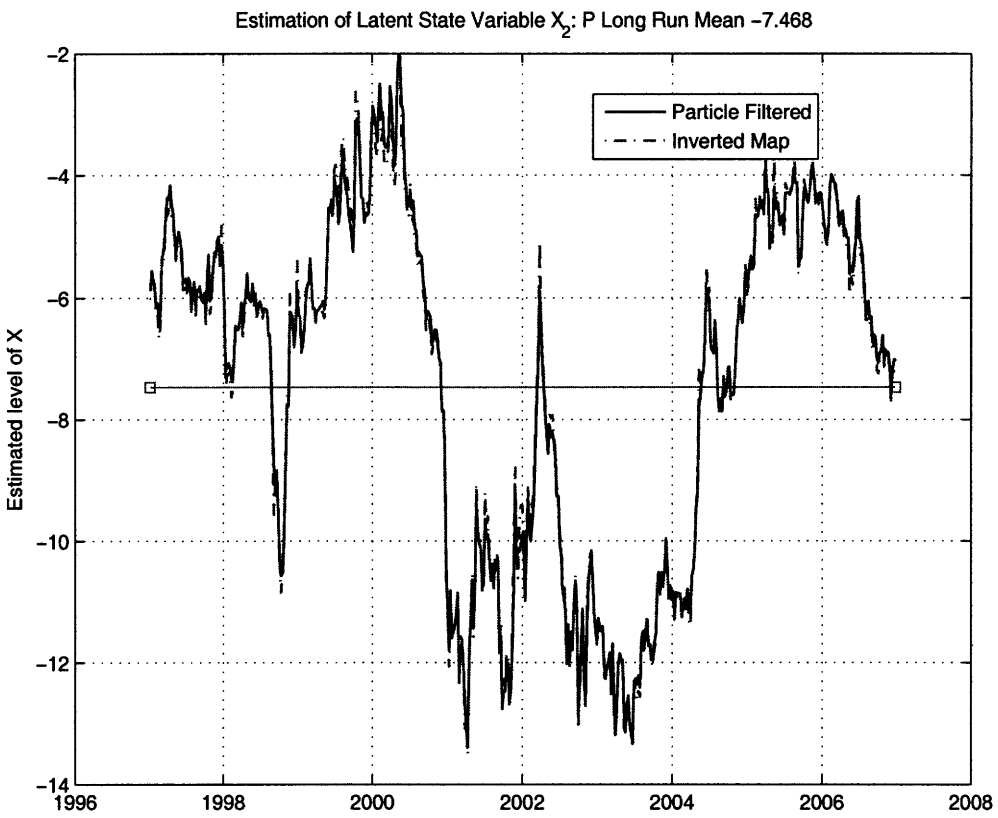

Figure 4-2: Estimated values of the latent state variable X_t^2. Horizontal bar fixed at the P measurable long run mean of X_t^2.

	Q Long Run Mean	P Long Run Mean
X^1	263.1865	261.7025
X^2	6.3522	-7.4678
X^3	281.6463	222.3386
θ	22.4568	4.1526

Table 4.3: Long Run Mean of X_t under the Q pricing measure, and the historical measure P.

Figure 4-4 shows the filtered values of the latent short rate $r(\tilde{X}_t)$. Unlike in other models the short rate is consistent with other short dated yields, e.g. the correlation between the short rate and six month LIBOR is .9879. As [51] highlights, this characteristic is strongly connected to the model's success in matching short dated maturities.

We highlight the fact that the long run mean of θ_t under Q is quite large, estimated

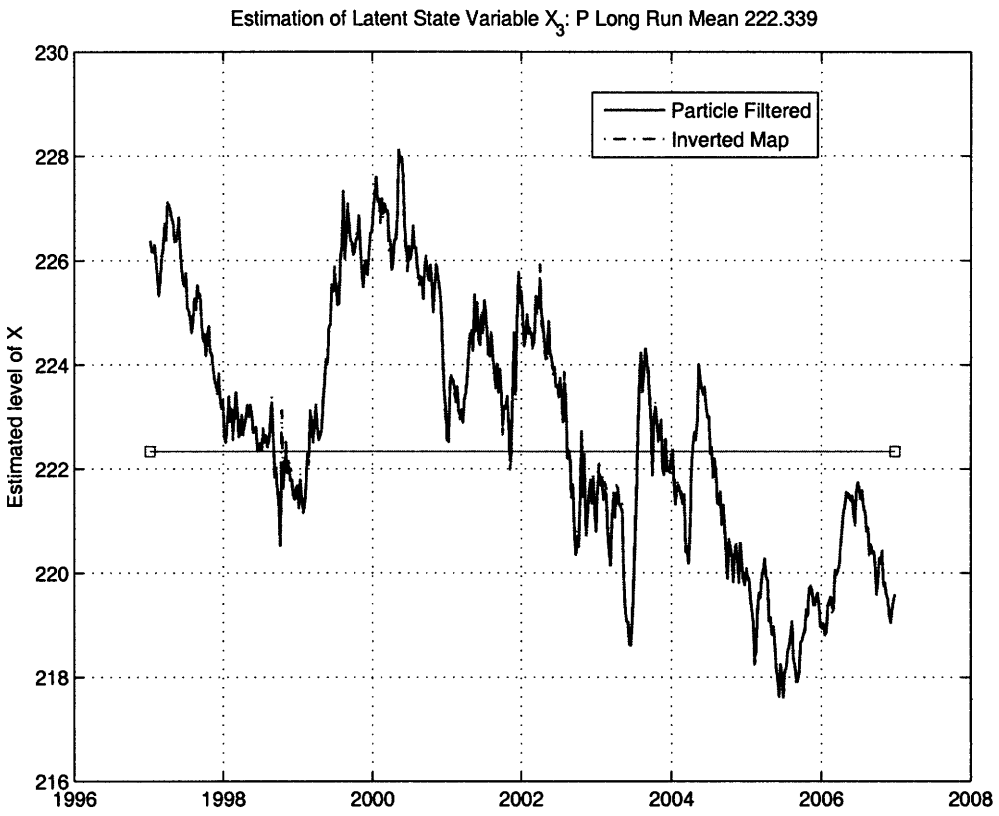

Figure 4-3: Estimated values of the latent state variable X_t^3. Horizontal bar fixed at the P measurable long run mean of X_t^3.

at 22.46 via Monte Carlo. Finally we note figure 4-5 confirms the long run mean of the stochastic intensity is zero under both measures, which is required if θ_t is to be bounded.

Moment Matching

Another measure of the model's ability to fit observed data is matching empirical moments. We compute unconditional moments of yields via Monte Carlo Integration, using the observed FOMC meeting schedule from our sample. A draw from the unconditional joint distribution of $\tilde{X}_t$ is obtained by simulating the joint dynamics via an Euler discritization for T years, such that the initial condition has no effect on the final state. Simulating the joint dynamics via an Euler discritization is detailed in Appendix B.1. To construct unconditional yield moments we take N draws from

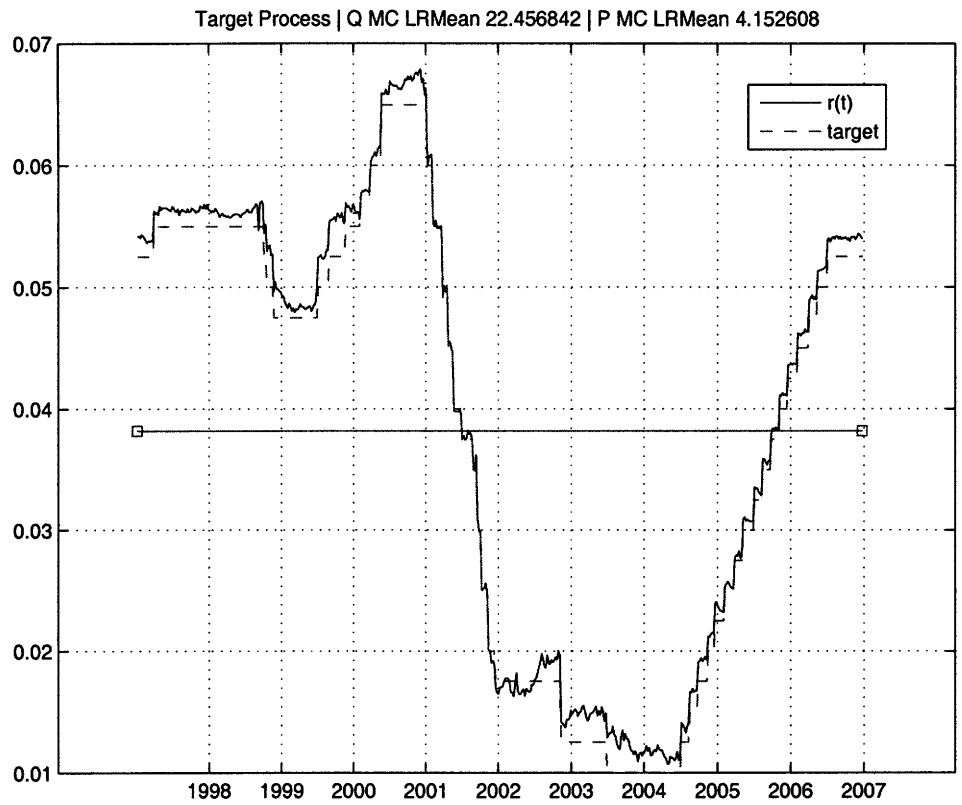

Figure 4-4: Estimated values of the latent short rate $r(\tilde{X}_t)$, and the observable Fed target rate.

the unconditional distribution and simulate each draw through the FOMC meeting schedule. We then compute model yields at weekly intervals to match the time grid of observed data. Sample moments are then easily computed from the generated model yields.

Figure 4-6-4-9 compare the model unconditional moments to sample estimates from the historical data. Figure 4-6 shows a good fit of model yield means to the sample mean of observed data. As seen in figure 4-7 the general shape of unconditional volatility matches the data well, however there does exist a noticeable spread between the two metrics. The model based estimate displays a hump at three years, which is related to pricing vectors as seen in figure 4-17 and 4-18. Unfortunately the model does not appear to possess any non-zero skew. The most encouraging result in unconditional moments focuses on the excess kurtosis estimate of figure 4-9.

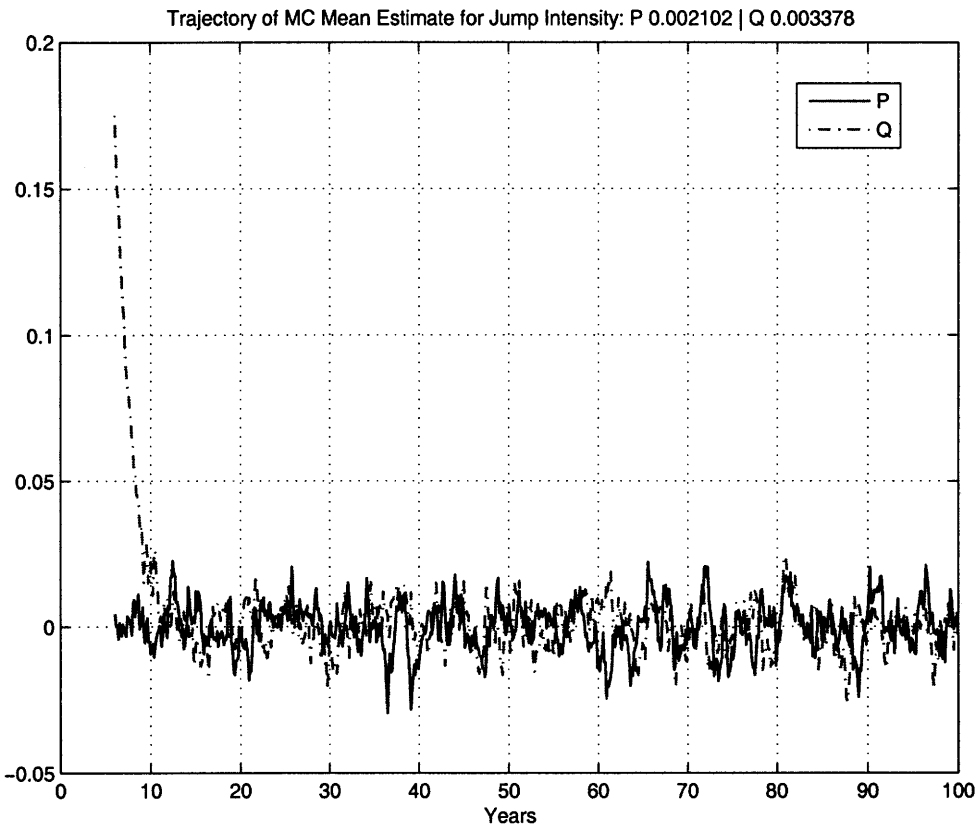

Figure 4-5: Monte Carlo verification of long run mean of jump intensity.

The existence of jumps in yield dynamics allows the model to match the non-normal behavior, especially at the short end of the yield curve.

Figure 4-10 displays the unconditional distribution of the short rate, which by construction is an affine function of the state variables. The figure highlights the small, though non-zero, probability of the short rate becoming negative. The negative short rate is due to the lack of a reflecting boundary at zero. Negative interest rates of any maturity violate no arbitrage pricing arguments, however in rare period short dated yields have been slightly negative. Typically the identified driver for negative yields is a flight to quality.

Tracking the Fed Target

Figure 4-11 plots the model's in-sample forecasting ability with respect to target changes during scheduled FOMC meetings. Estimates are constructed by generating

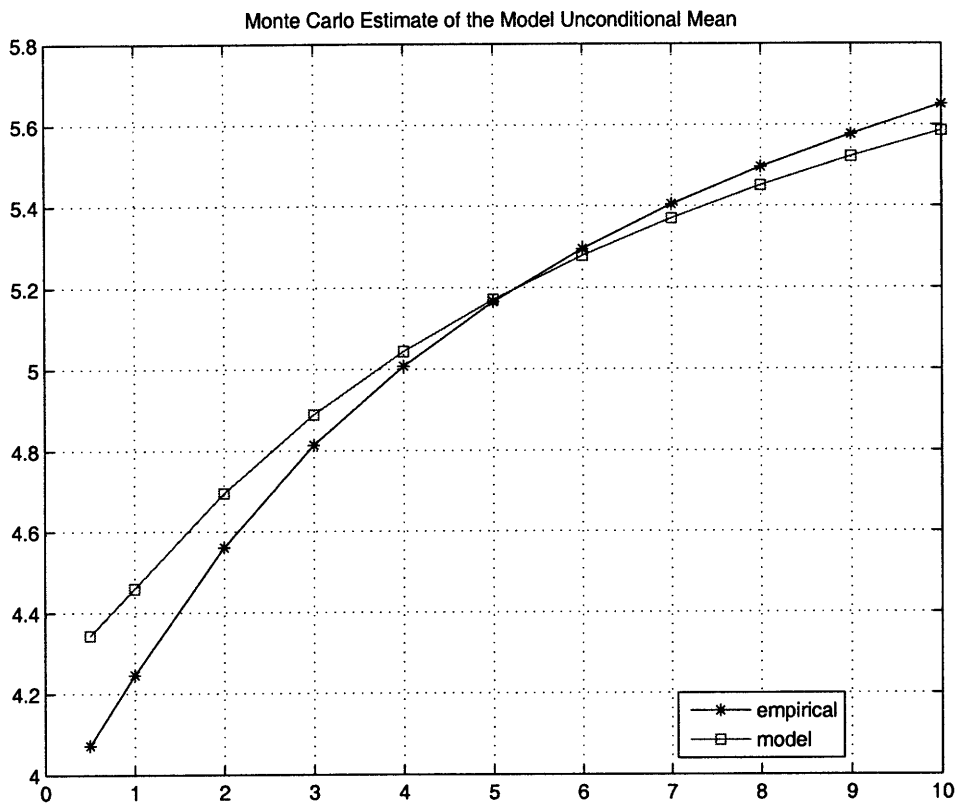

Figure 4-6: Monte Carlo estimate of model yields unconditional mean.

conditional distributions of θ_t at the close of each scheduled FOMC meeting. The distributions are obtained from Monte Carlo simulations using the optimal model parameters in table 4.1. The non-linear switching function in the dynamics of $d\theta_t$, results in a two sided Poisson as a conditional distribution. Where each arrival of the Poisson is a 25 basis point jump in θ. The maximum likelihood estimate (MLE) is then the 25 basis point value with the maximum number of occurrences under the historical measure. Finally $E[\Delta\theta]$ is easily found by integrating the Monte Carlo draws.

Over the ten year plus sample, the model displays an impressive ability to predict changes in the target rate. The model predicts the correct sign for all FOMC target changes, including all no-change meetings. Out of the 80 scheduled meetings in the sample, the model predicts 63 with absolute precision. Table 4.4 documents the in-sample prediction errors. The absolute value of the maximum error is 25 basis points, meaning at most the model misses a single Poisson jump. The model appears to have

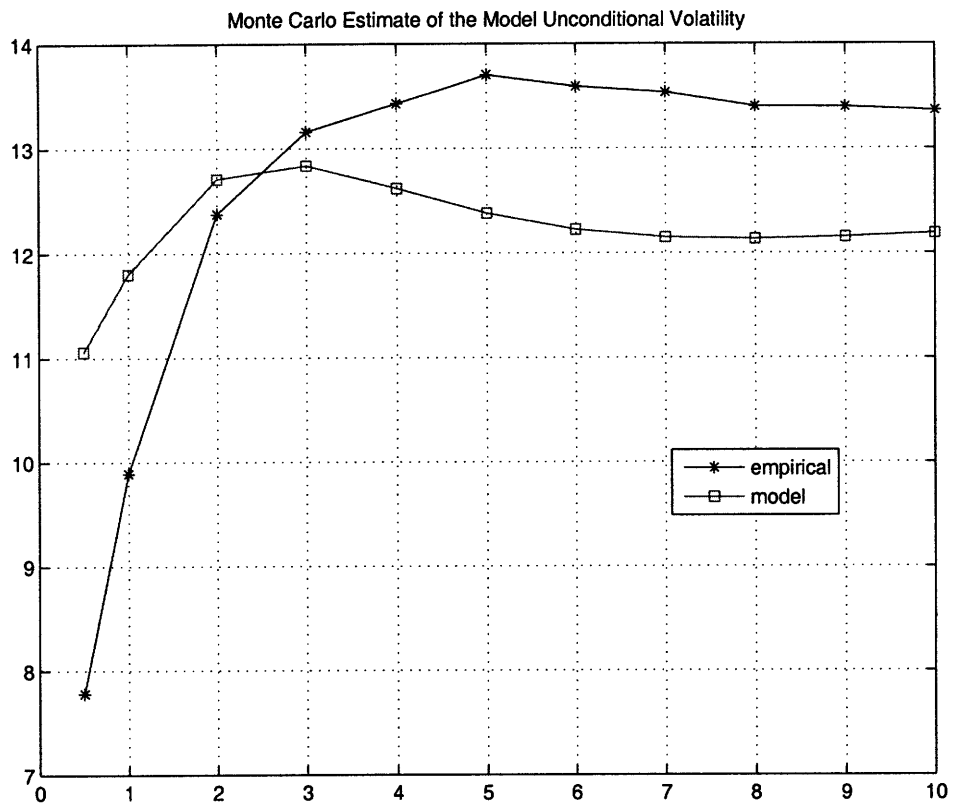

Figure 4-7: Monte Carlo estimate of model yields unconditional volatility.

the most difficulty with the tightening period which began on June 30th 2004, and resulted in 17 consecutive 25 basis point increases. Over this time period the mean of the conditional distribution is always within a few basis points of the realized target change. However in several cases the maximum point density lies at zero, hence our estimator predicts the incorrect jump size.

From table 4.1 we find the time t scheduled meeting jump intensity kernel as

$$\lambda_t = \lambda_o + \lambda_\theta \theta_t + \lambda_1 X_t^1 + \lambda_2 X_t^2 + \lambda_3 X_t^3 \tag{4.2}$$

$$= 0.51 - 9,941.19\theta_t + 5.18X_t^1 + 141.16X_t^2 - 0.01X_t^3 \tag{4.3}$$

where the actual jump intensity is the absolute value of λ_t. An alternative view is to

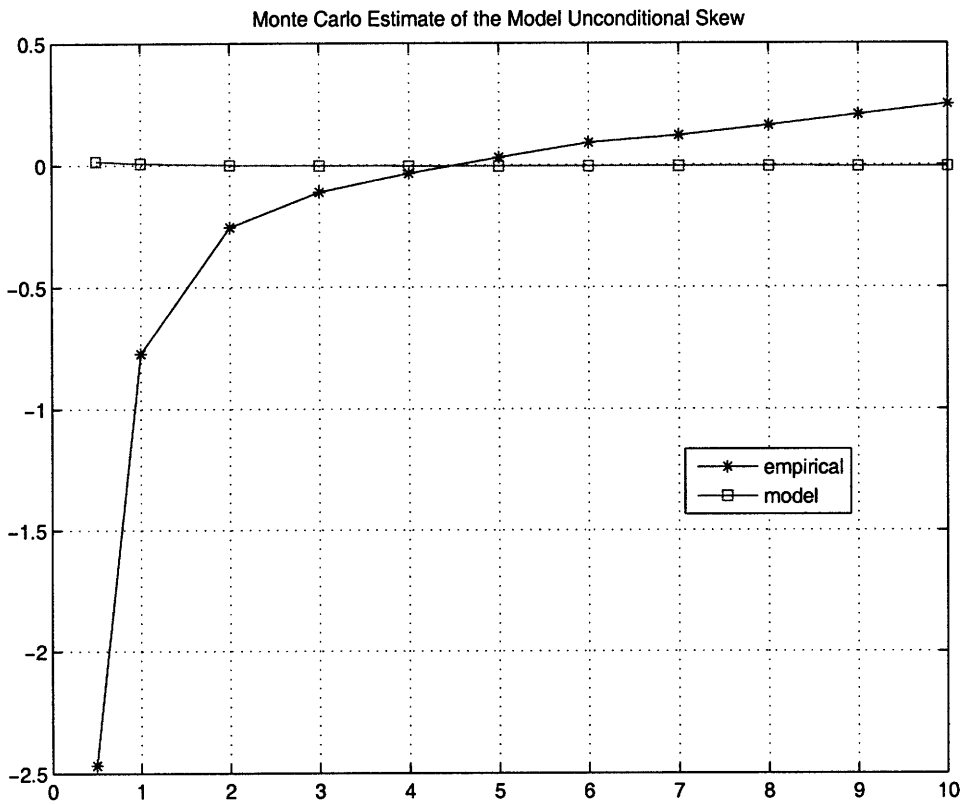

Figure 4-8: Monte Carlo estimate of model yields unconditional skew.

transform eq(4.3) into deviations from long run means.

$$\lambda_t = 0.51 - 9,941.19\theta_t + 5.18X_t^1 + 141.16X_t^2 - 0.01X_t^3 \quad (4.4)$$

$$= 0.2028 - 9,941.19(\theta_t - \overline{\theta}) \quad (4.5)$$

$$+ \; 5.18(X_t^1 - \overline{X}_1) + 141.16(X_t^2 - \overline{X}_2) - 0.01(X_t^3 - \overline{X}_3)$$

In steady state the 0.2028 constant term results in an average of $[(0.2028 * 8)/365] = 0.0044$ jumps per calender year. This is roughly one jump every 225 years, which is consistent with the view of a stable system. Our estimate of λ_θ is very close to the $-9,408.9$ estimate in [51]. It follows that a single 25 bp jump at the current meeting will push the expected change at the next meeting downward by $9408.9 * .0025/365 * 25 = 1.61$ basis points.

Figure 4-12 shows a decomposition of the future expected target change over the next scheduled FOMC meeting. This is approximately equal to $\lambda_t h$, where h is the

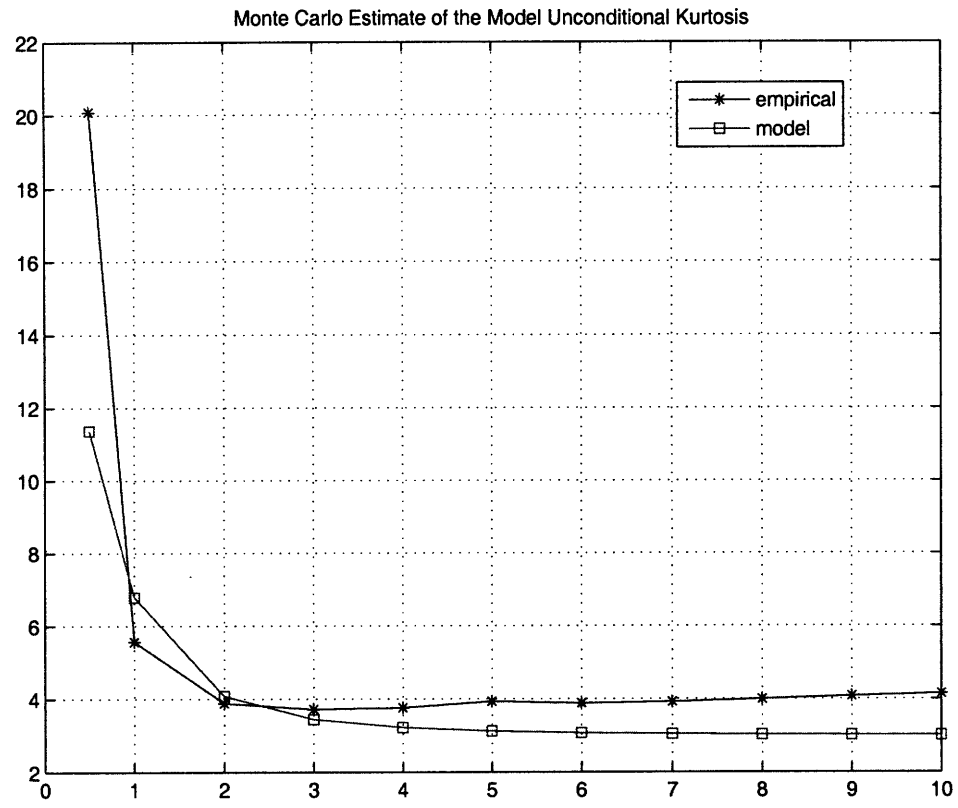

Figure 4-9: Monte Carlo estimate of model yields unconditional kurtosis.

length of a FOMC meeting. Where the approximation is in holding λ_t constant during the meeting. The figure provides graphical support to the model parameters associated with the jump process given in table 4.1. From 4-12 we find the target itself and X_t^2 predominately shape monetary policy.

It is also of interest to infer which yields drive changes in the target. Inverting the pricing equation for the three yields observed without error would provide a weighting of yields with respect to λ_t. However for scheduled meetings far in the future this would be inaccurate. Additionally this limits by construction which yields influence the jump process. Instead we take an imperial approach, and calculate the correlation of yields and the time series of expected target changes at the next scheduled FOMC meeting. Figure 4-13 plots the correlation coefficients from this calculation. We find the one and two year yields to have the most influence on monetary policy.

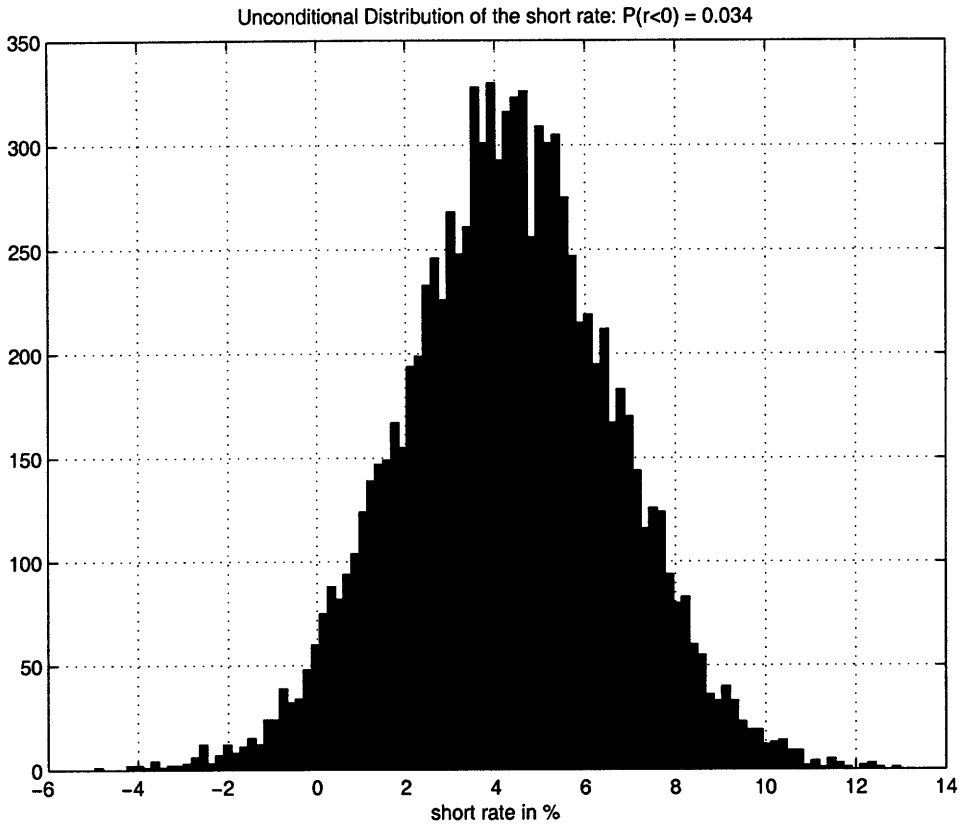

Figure 4-10: Monte Carlo draws of the short rate as a histogram.

Model Extensions: θ Feedback

From the model construction of section 3.2.1 we note there exists a one way influence between X_t and θ_t. That is X_t appears in the drift of θ_t, but the channel is not open in the other direction. We explore allowing a feedback from θ_t to the drift of X_t. Recall the dynamics of X_t under the historical measure.

$$dX_t = \mu^P(X_t)dt + \sigma(X_t)dW_t^P \tag{4.6}$$

the extension in question defines $\mu^P(X_t)$ in terms of $\tilde{X}_t$. That is

$$\mu^P = K_o^P + K_1^P \cdot [X_t, \theta_t]^\top = \begin{bmatrix} k_{01}^P \\ k_{02}^P \\ k_{03}^P \end{bmatrix} + \begin{bmatrix} k_{11}^P & 0 & 0 & 0 \\ k_{21}^P & k_{22}^P & k_{23}^P & k_{24}^P \\ k_{31}^P & k_{32}^P & k_{33}^P & k_{34}^P \end{bmatrix} \cdot [X_t, \theta_t]^\top \tag{4.7}$$

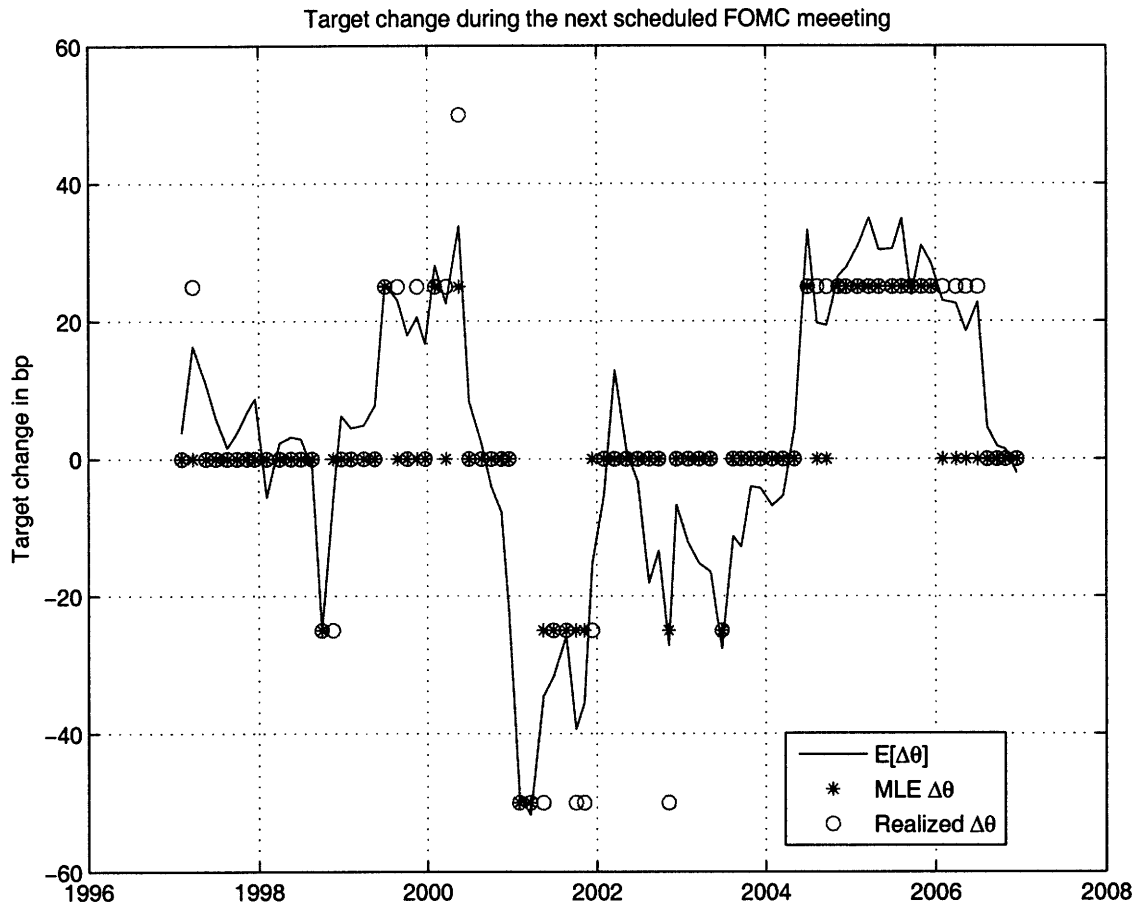

Figure 4-11: Short dated model predictions of target changes at the next FOMC meeting. Estimates are maximum likelihood values via a Monte Carlo generated distribution of θ.

where we set $k_{14} = 0$ to maintain admissibility in light of the target's non-positive attainment. We follow the flexible risk specification of section 3.2.3 to allow a similar feedback under Q.

$$\mu^Q = K_o^Q + K_1^Q \cdot [X_t, \theta_t]^\top = \begin{bmatrix} k_{01}^Q \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} k_{11}^Q & 0 & 0 & 0 \\ k_{21}^Q & k_{22}^Q & k_{23}^Q & k_{24}^Q \\ k_{31}^Q & k_{32}^Q & k_{33}^Q & k_{34}^Q \end{bmatrix} \cdot [X_t, \theta_t]^\top \quad (4.8)$$

Point estimates for this extension are found in table 4.5. In general we find very similar estimates as in table 4.1, with a few exceptions. Though not statically significant at the 5% level, λ_2 and k_{24}^P do have a noticeable impact on the dynamics.

FOMC Date	Realized $\Delta\theta$ (bp)	Predicted $\Delta\theta$ (bp)
26-Mar-97	25	0
18-Nov-98	-25	0
25-Aug-99	25	0
17-Nov-99	25	0
22-Mar-00	25	0
17-May-00	50	25
16-May-01	-50	-25
03-Oct-01	-50	-25
07-Nov-01	-50	-25
12-Dec-01	-25	0
07-Nov-02	-50	-25
11-Aug-04	25	0
22-Sep-04	25	0
01-Feb-06	25	0
29-Mar-06	25	0
11-May-06	25	0
30-Jun-06	25	0

Table 4.4: In-sample prediction errors of near dated target changes during scheduled FOMC meetings. Maximum absolute prediction error is 25 basis points.

This is not very surprising, considering X_t^2 and θ_t are the primary drivers for target jumps. Though this does alter some of the performance metrics discussed above, it does so in a slight fashion. In the interest of building a parsimonious model we continue with the original dynamics given in section 3.2.1.

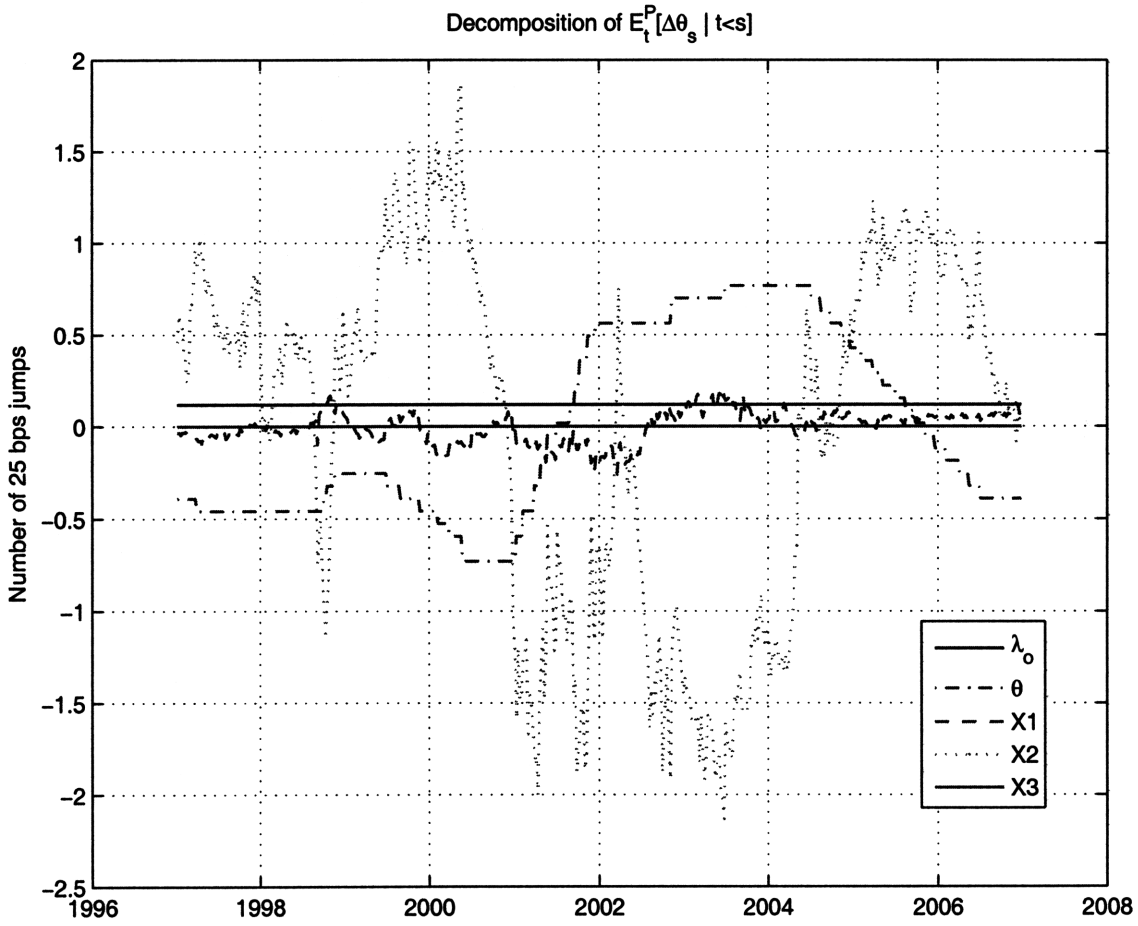

Figure 4-12: Decomposition of future expected target changes over the next scheduled FOMC meeting, where the y-axis is in units of 25 basis points. Future expected target changes are approximately equal to $E_t[\lambda_s | s \geq t]h$. Where the next scheduled meeting is at time s and h is the length of the meeting.

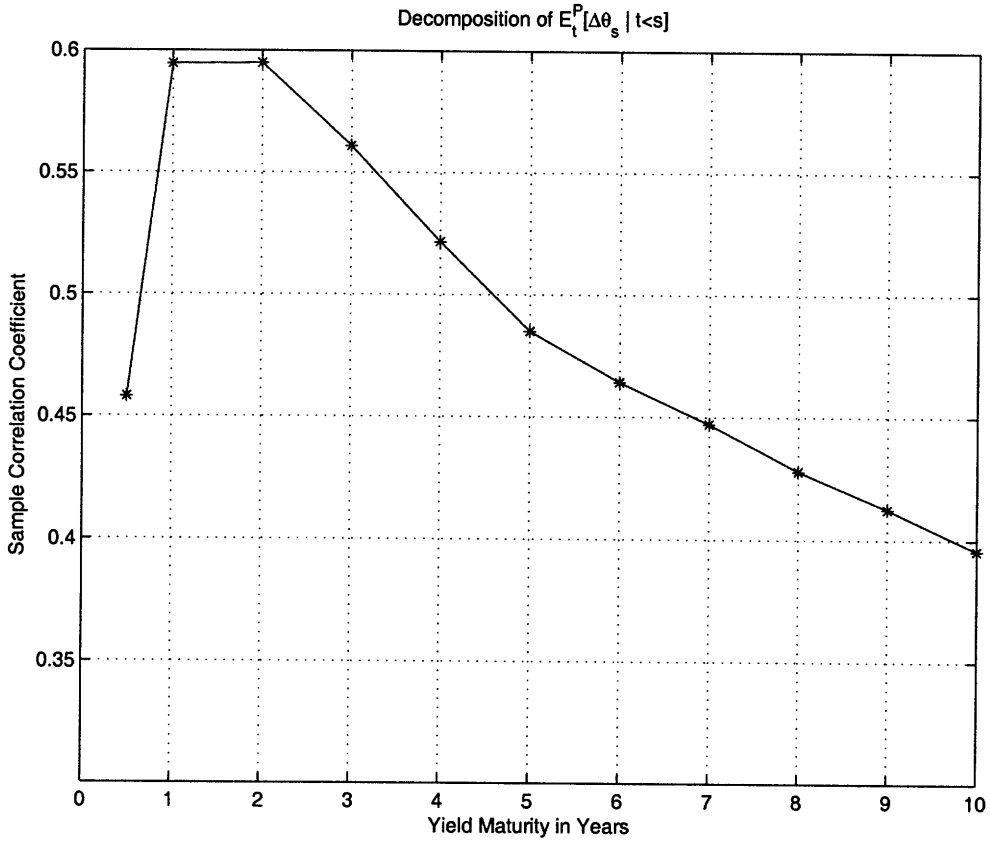

Figure 4-13: Decomposition of future expected target changes over the next scheduled FOMC meeting with respect to current yields. Future expected target changes are approximately equal to $E_t[\lambda_s | s \geq t]h$. Where the next scheduled meeting is at time s and h is the length of the meeting. Correlation is the standard sample correlation on first differences.

Parameter	Estimate	Standard Error	t-Statistic
k_{o1}^{Q}	1065.5036	1847.5416	0.5767
k_{11}^{Q}	1.3225	0.0557	**23.7578**
k_{21}^{Q}	0.0678	0.0632	1.0733
k_{22}^{Q}	2.2202	0.0946	**23.4610**
k_{23}^{Q}	-0.0968	2.8833	-0.0336
k_{24}^{Q}	0.0105	2.0005	0.0052
k_{31}^{Q}	-0.0443*1e-6	0.0000	-0.0024
k_{32}^{Q}	0.1625	4.8418	0.0336
k_{33}^{Q}	0.00868*1e6	0.0000	0.0026
k_{34}^{Q}	0.0087	0.9000	0.0096
b_{21}	0.0052*1e-3	0.0014	0.0039
b_{31}	0.0220	1.3800	0.0160
ρ_o	-0.1782	0.1715	-1.0392
ρ_1	0.0999*1e-3	0.0001	1.1551
ρ_2	0.0048*1e-6	0.0000	0.0031
ρ_3	0.2046*1e-3	0.0061	0.0336
λ_o	6.8417	1851.9019	0.0037
λ_θ	-9926.6075	628.6534	**-15.7903**
λ_1	2.8542	3.0605	0.9326
λ_2	509.4167	278.4925	1.8292
λ_3	-0.0015	0.2953	-0.0049
k_{o1}^{P}	2513.5931	4427.9416	0.5677
k_{o2}^{P}	-0.0110	97.7537	-0.0001
k_{o3}^{P}	897.8020	26886.3990	0.0334
k_{11}^{P}	3.1304	0.8602	**3.6393**
k_{21}^{P}	0.0133	0.0350	0.3813
k_{22}^{P}	0.7094	0.5212	1.3612
k_{23}^{P}	-0.0181	0.5785	-0.0312
k_{24}^{P}	18.2287	25.3658	0.7186
k_{31}^{P}	0.2377	7.0922	0.0335
k_{32}^{P}	0.4712	14.2600	0.0330
k_{33}^{P}	1.4624	0.6662	**2.1952**
k_{34}^{P}	0.0203	101.1477	0.0002

Table 4.5: Point parameter estimates. Estimates are calculated via the simulated maximum likelihood technique described in section 3.3.1. Sample period is from January 1997 to January 2007, including 521 weekly samples. Standard errors are computed via the product of outer gradients [3].

Model Extensions: θ Jump Risk

Section 3.2.3 presented a number of risk premium specifications for the jump process. For clarity we stress that securities linked to the target rate will possess a risk premium due to the stochastic intensity. The jump risk we explore in this section concerns the risk premium for jumps conditioned on $\tilde{X}_t$. From a market perspective this is closely linked to the premium one would demand for buying a Fed future contract immediately before the scheduled FOMC meeting. With this distinction in mind we report on the two main jump risk specifications of section 3.2.3.

The two specifications explored are

$$\lambda_t^Q = \lambda_q \left(\lambda_o + \lambda_\theta \theta_t + \lambda_X X_t\right) \tag{4.9}$$

$$\lambda_t^Q = \lambda_o^Q + \lambda_\theta^Q \theta_t + \lambda_X^Q X_t \tag{4.10}$$

We estimated separate models using the above specification. The resulting model parameters as well as previously reported performance metrics were unchanged. This is not surprising considering the very small sample which identifies jump risk. Of the 521 weekly samples, 80 contain scheduled FOMC meetings. Unfortunately with the small sample we are unable to distinguish risk associated from X_t and risk defined via equations 4.9 and 4.10.

4.2 Pricing Errors

Our estimation scheme results in uniformly low *in-sample* errors on all observed yields. Application of a Bayesian Particle Filter supports these results, and shows them to be robust when assuming all observations possess additive error. Information contained in Fed Future contracts are not used during the estimation process, and as such are viewed as *out-of-sample* errors. In strong support of the model out-of-sample errors are consistent with the low in sample errors. We present out-of-sample errors in a contract pricing framework, as well as in the context of market implied target changes.

Yields

Table 4.6 reports the in-sample pricing errors at the optimal model parameters of table 4.1. Recall our simulated maximum likelihood scheme observed three yields without error, enabling us to invert the pricing equation to infer the latent state space. The remaining yields were assumed to possess additive error of the form

$$Y(t,T) = C_o(t,T) + C_{\tilde{X}}(t,T)\tilde{X}_t + \epsilon_t \tag{4.11}$$

where ϵ_t is a eight dimensional zero mean multivariate Gaussian noise term with a diagonal covariance matrix of equal entries. Under this framework a single parameter controls the full distribution of ϵ_t. We choose the standard deviation as the control parameter, and denote it as σ_ϵ. During the SML estimation scheme this parameter is easily integrated out during each evaluation of the likelihood function. Under the optimal model parameters of table 4.1, we estimate $\sigma_\epsilon = 3.5018$. We also calculate the root mean squared error (RMSE) of the in-sample pricing errors . The first column of table 4.6 shows in-sample pricing errors to be below five basis points, which is well within the bid ask spread of the associated swap contracts.

Column two of table 4.6 reveals the RMSE when the Bayesian Particle Filter of section 3.3.2 is used to infer the latent states of X_t. As with column one, the Particle Filter reveals an excellent ability to match observed yields. This is especially true at the short end of the yield curve, which is traditionally a challenge for DTSM models

[52]. As discussed in [51], the target may be viewed as a proxy for very short dated maturity yields. Indeed the target is constructed as the Fed's goal for the overnight interest rate, as such it can be viewed as a stable or smoothed proxy for short dated yields.

Maturity (yr)	RMSE via map (bp)	RMSE via Particle Filter (bp)
0.5	0	3.3013
1.0	4.8987	3.5357
2.0	0	2.4043
3.0	3.1408	1.4717
4.0	4.2056	1.4582
5.0	4.2991	1.7124
6.0	3.7302	1.4562
7.0	2.9379	1.0130
8.0	2.0480	0.7599
9.0	1.1499	1.1895
10.0	0	1.8676

Table 4.6: In sample pricing errors of yields in basis points. RMSE via map is found by observing the six month, two year, and ten year yields with no error and inverting the measurement equation. RMSE via Particle Filter is obtained by applying the Bayesian Particle Filter of section 3.3.2.

Near Dated Fed Future Contracts

We choose to treat the Fed future contracts of section 3.4.3 as out-of-sample observations, providing an additional test for model performance. Fed future contracts settle at the end of each month, and are priced at 100 minus the arithmetic average of the Fed Funds rate during the contract month. For ease of notation we normalize contract prices and deal directly with the arithmetic average. The time t price for a normalized contract which settles in month s, is given by

$$f_t^s = \frac{1}{m}\left[\sum_{i=1}^{t}\tilde{r}_i + \sum_{i=t+1}^{m} E_t(\tilde{r}_i)\right] \tag{4.12}$$

where m is the number of days in month s and $\tilde{r}_i$ is the Fed funds rate on day i. For clarity we note the expectation is with respect to Q, the pricing measure. To

construct prices under our model framework we transform eq(4.12) to be a function of the target rate

$$
\begin{aligned}
f_t^s &= \frac{1}{m}\left[\sum_{i=1}^{t}\tilde{r}_i + \sum_{i=t+1}^{m} E_t(\tilde{r}_i)\right] && (4.13)\\
&= \frac{t}{m}\overline{r}_{i\leq t} + \frac{m-t}{m}E_t[\overline{r}_{i>t}] && (4.14)\\
&= \frac{t}{m}(\theta_{i\leq t} + \overline{\eta}_{i\leq t}) + \frac{m-t}{m}E_t[(\theta_{i>t} + \overline{\eta}_{i>t})] && (4.15)\\
&\approx \frac{t}{m}(\theta_{i\leq t}) + \frac{m-t}{m}E_t[\theta_{i>t}] && (4.16)
\end{aligned}
$$

where $\overline{r}$ is the sample average of the funds rate, θ is the target rate, and $\overline{\eta}$ is the sample average of the associated Fed tracking error. As noted in section 3.2 the tracking error on average is very low, on the order of basis points, allowing us to safely ignore this term.

Under expectation the target only changes during scheduled FOMC meetings. Within our model we can easily see this via the dynamics for the target outside of schedule meetings as given in section 3.2.2

$$
d\theta_t = J_\theta\left(dN_t^u - dN_t^d\right) \quad \text{for t} \notin \text{scheduled FOMC meeting} \tag{4.17}
$$

where the jump intensity of dN_t^u and dN_t^d is equal to $\overline{\lambda}$, and J_θ equals 25 basis points. Since the expectation of eq(4.16) is a linear function of θ, the two equal jump intensities will cancel out under expectation. Similarly one could construct a compensated process and note the drift is zero.

Building on this insight we construct a rolling time series of Fed future contracts which are exposed to potential target changes during scheduled meetings. Which contract we select depends on the position of future scheduled FOMC meetings. Consider the case when the next scheduled FOMC meeting is on day j in the current month s, eq(4.16) then becomes

$$
f_t^s = \frac{j}{m}\theta_t + \frac{m-j}{m}E_t[\theta_{j+}] \tag{4.18}
$$

where the notation assumes there were no unscheduled meetings in the month from $[0,t]$, and $t < j \leq m$. We indicate the beginning of a scheduled meeting with the index j, and the close of a meeting by $j+$. It follows that θ_{j+} is the target rate at the close of the meeting. Since under expectation the target does not change outside of scheduled meetings, eq(4.18) has the same form when the next meeting is in month $s+1$. As $j \to m$ the influence of FOMC meetings on future contracts falls to zero. When the next meeting is in month $s+2$, an alternative to f_t^s exists which possesses a larger derivative with respect to $E_t[\theta_{j+}]$. Consider the case when the next scheduled FOMC meeting is on day j in the current month s, however the following FOMC meeting is in month $u \geq s+2$. In this case the rolling time series chooses f_t^{s+1}, where

$$f_t^{s+1} = \frac{1}{m}\sum_{i=1}^{m} E_t[\theta_i] \tag{4.19}$$

$$= E_t[\theta_{j+}] \tag{4.20}$$

A flexible means to compute $E_t[\theta_{j+}]$ is Monte Carlo integration under Q dynamics. Alternatively an analytical solution is available by exploiting the linearity of expectation. Write θ_t as a compensated Poisson process

$$d\theta_t = (\lambda_o + \lambda_\theta \theta_t + \lambda_X^\top X_t)dt + dM_t^Q \tag{4.21}$$

where M_t^Q possesses a non-Gaussian zero mean distribution. Since the drift of $d\theta_t$ is linear in $\tilde{X}$ we can express the conditional expectation of $\tilde{X}$ at the close of the FOMC meeting as

$$E[\tilde{X}_{j+}|\tilde{X}_j] = a + B\tilde{X}_t \tag{4.22}$$

$$= (I - \exp(-Kh))K^{-1}K_o + \exp(-Kh)\tilde{X}_j \tag{4.23}$$

where j indicates the start of a meeting, $j+$ the close, h the length of a meeting, and

$$K_o = \begin{bmatrix} \lambda_o \\ k_{01}^Q \\ k_{02}^Q \\ k_{03}^Q \end{bmatrix} \tag{4.24}$$

$$K = \begin{bmatrix} -\lambda_\theta & -\lambda_1 & -\lambda_2 & -\lambda_3 \\ 0 & k_{11}^Q & 0 & 0 \\ 0 & k_{21}^Q & k_{22}^Q & k_{23}^Q \\ 0 & k_{31}^Q & k_{32}^Q q & k_{33}^Q \end{bmatrix} \tag{4.25}$$

To find $E[\tilde{X}_j|\tilde{X}_t]$ we again leverage the linearity of conditional expectation in affine processes, and note $E_t[\theta_j] = \theta_t$.

$$\begin{aligned} E[X_j|X_t] &= c + DX_t & (4.26) \\ &= (I - \exp(-K(j-t)))K^{-1}K_o + \exp(-K(j-t))X_t & (4.27) \end{aligned}$$

where

$$K_o = \begin{bmatrix} k_{01}^Q \\ k_{02}^Q \\ k_{03}^Q \end{bmatrix} \tag{4.28}$$

$$K = \begin{bmatrix} k_{11}^Q & 0 & 0 \\ k_{21}^Q & k_{22}^Q & k_{23}^Q \\ k_{31}^Q & k_{32}^Q q & k_{33}^Q \end{bmatrix} \tag{4.29}$$

Figure 4.2 shows the out-of-sample pricing errors of the constructed rolling time series of Fed future contracts. The root means squared error over the sample is 5.090

bp. The model's ability to match Fed future contracts provides strong support to the flexible jump dynamics. Since the future contracts are out-of-sample, this result also provides substantial evidence that information in the yield market is also contained in the Fed future market.

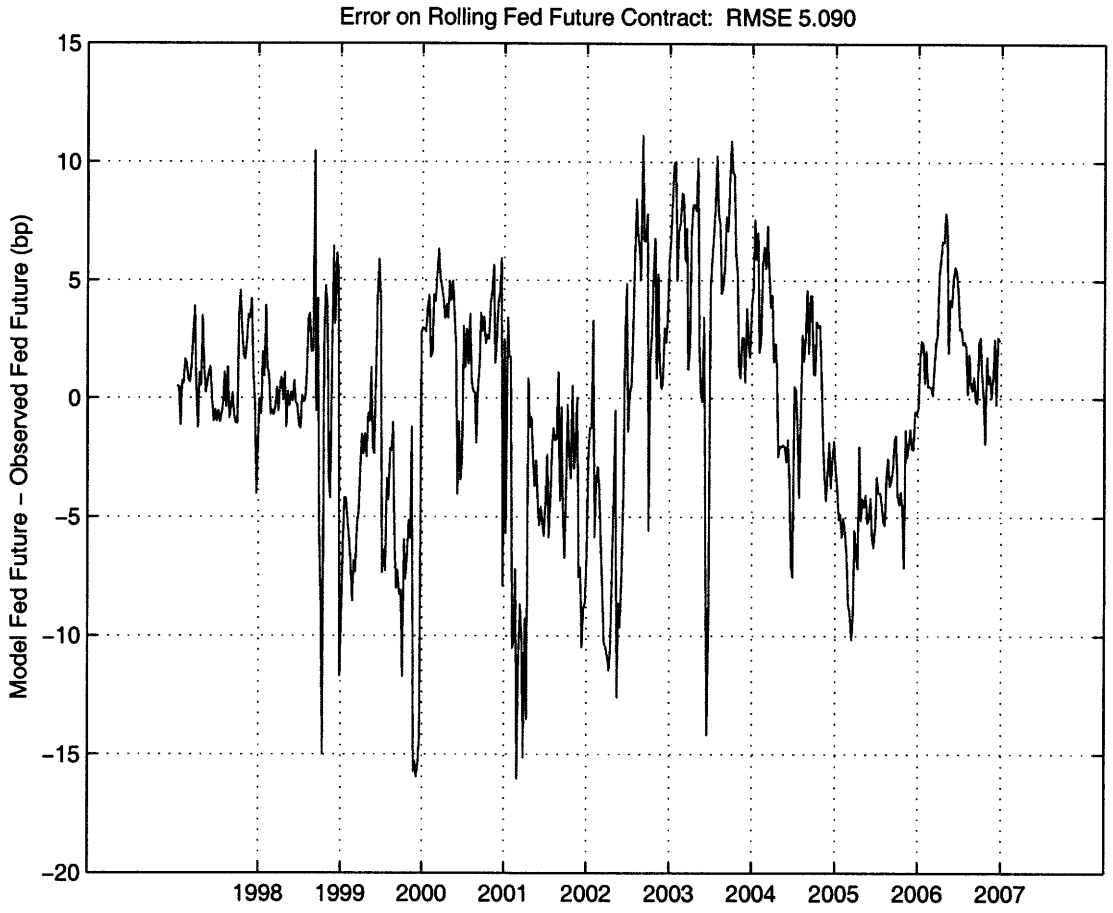

Figure 4-14: A time series of out-of-sample pricing errors for near dated Fed Future contracts.

An alternative view of the model's ability to match information contained in the Fed future contracts, is to invert the pricing relation to reveal jump intensities. Specifically we invert eq(4.18) and eq(4.20) to expose the Fed future implied expected change in the target at the next scheduled FOMC meeting. We can then easily compare this time series to the equivalent metric in our model. Figure 4.2 plots the two time series for comparison.

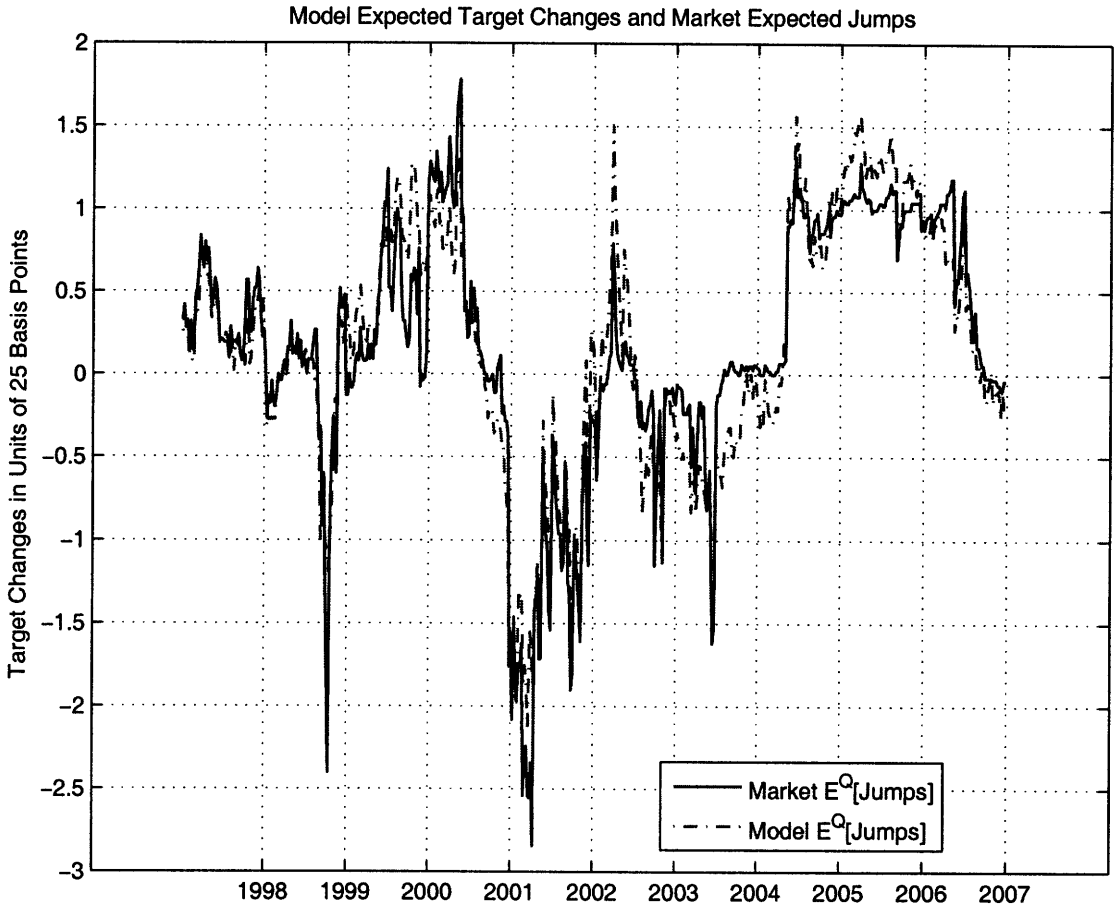

Figure 4-15: Time series of the number of 25 basis point jumps in the Fed target rate expected to occur at the next scheduled FOMC meeting. Model E^Q[Jumps] are taken from the model using the optimal model parameters of table 4.1. Market E^Q[Jumps] are obtained by inverting the pricing relation for Fed future contracts.

4.3 Yield Response to Shocks

Figure 4-16 plots the loadings of yields on the first three principle components (PC). Since [46] first reported on the ability of three PCs to explain 99.95% of the variation in yields, the literature has been referring to the three components as level, slope, and curvature factors. Due to the shape of the curves in figure 4-16, we label PC1 as the level factor, PC2 as the slope factor, and PC3 as the curvature factor.

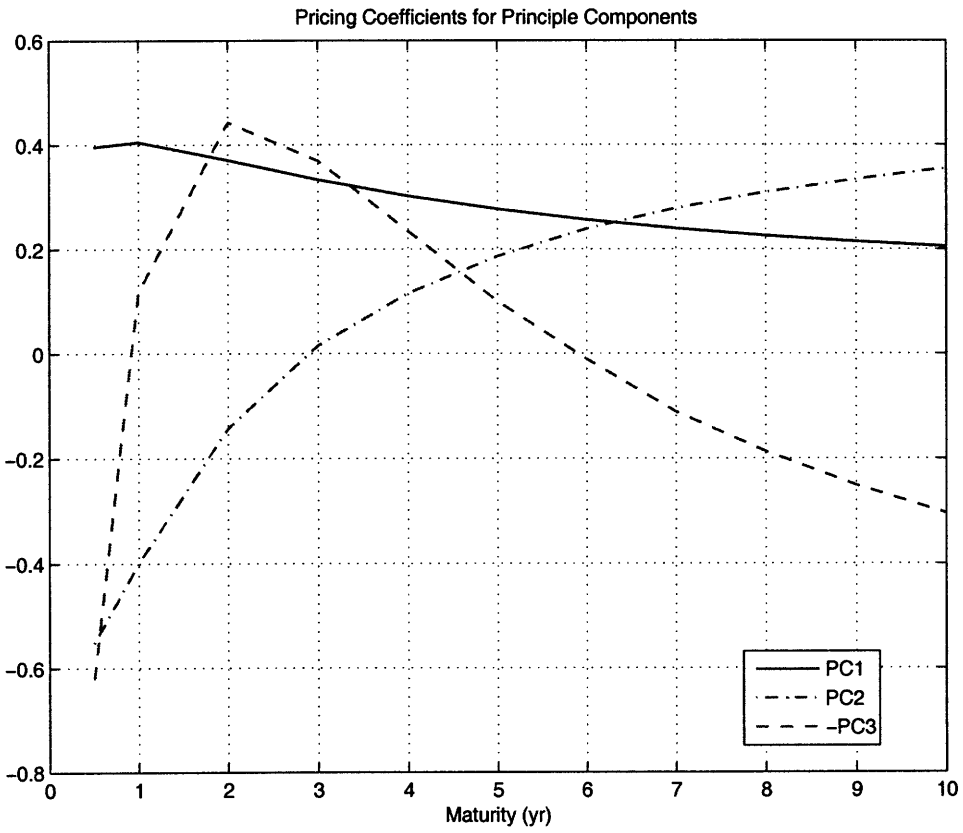

Figure 4-16: Loadings for the first three Principle Components of Yields.

Figure 4-17 and 4-18 display normalized price loadings for model parameters given in section 4.1. Specifically we plot normalized $C_{\tilde{X}}$ of the pricing equation

$$Y(t,T) = C_o(t,T) + C_{\tilde{X}}(t,T)\tilde{X}_t + \epsilon_t \tag{4.30}$$

where $C_{\tilde{X}} = [C_\theta, C_1, C_2, C_3]$. To view on the same scale we normalize each coefficient to reflect the associated change in yields for a conditional one standard deviation

shock to X_t. C_θ is not normalized, thus it represents the yield response to a 1% change in the target. We identify C_2 as a curvature factor, as well as C_θ and C_3 as a slope factor. The remaining C_2 is seen as a hybrid of the traditional factor labels. It possesses traits of a slope factor at the mid to long end of the yield curve. However near short to mid maturities is displays a strong curvature effect.

The meeting schedule of the FOMC injects the time variation found in $C_{\tilde{X}}(t,T)$. The only dynamics to change with the arrival of a scheduled meeting are the dynamics associated with the target itself or $d\theta_t$. However the coupling via the jump intensity leads all state variables to display time variation, via the λ coefficients of eq(3.22). It follows that the time variation of $C_{\tilde{X}}$ is greatest across a scheduled meeting. Figure 4-17 captures $C_{\tilde{X}}$ directly before a meeting, while 4-18 plots $C_{\tilde{X}}$ immediately after a scheduled meeting. A convenient metric to view this seasonality is the days till the next scheduled FOMC meeting. As this metric increases the influence of λ_t and the associated jumps in θ have less of an effect on prices. Since X_3 doesn't effect the jump intensity of θ_t, it possess a very small amount of time variation. For similar reasons we find C_2 to display the most time variation, which is supported by X_2 dominating the jump intensity.

Monetary Policy Shocks

An active area of research is focused on identifying the response of asset prices to shocks in monetary policy. We define shocks in monetary policy as unanticipated changes in the Fed target rate. Since Fed future contracts are linked to the average target rate over a single month, they serve as a natural choice to measure shocks to the target rate. Our model construction provides a natural means to measure monetary policy shocks and the resulting effect on yields.

Section 4.2 presents a rolling time series of Fed future contracts, which are constructed for maximum exposure to the next scheduled FOMC meeting. Model prices of these contracts are formulated by computing the expected target change under the pricing measure. Shocks however are measured under the historical P measure. We define a shock to be the difference between the realized target change and the P

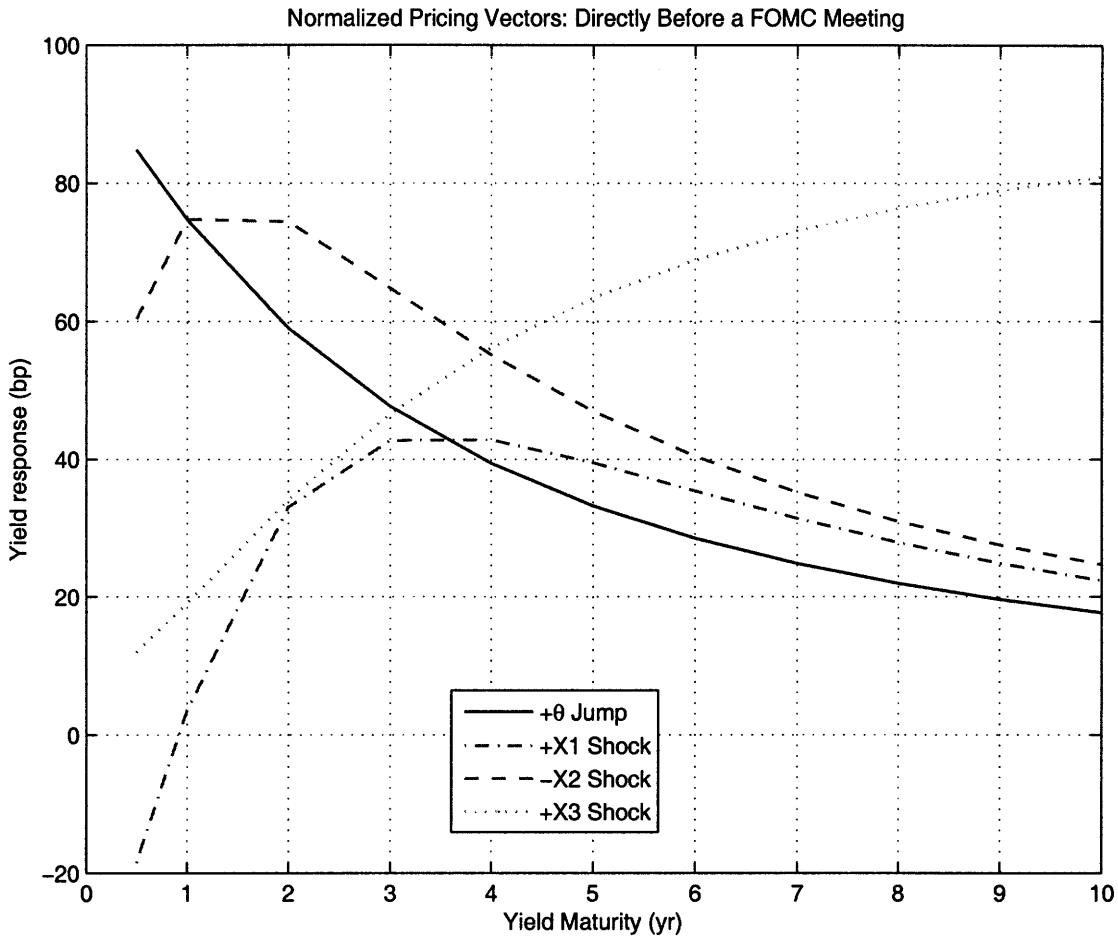

Figure 4-17: Normalized pricing vector. C_X is normalized to display the response to a 1 standard deviation shock. C_θ requires no such normalization.

measurable expected change. The challenge in using Fed future contracts to extract shocks is identifying the risk premium which distinguishes the two measures.

Recall the pricing relation for a time t Fed future contract given in section 4.2

$$f_t^s = \frac{j}{m}\theta_t + \frac{m-j}{m}E_t^Q[\theta_{j+}] \tag{4.31}$$

where there are m days in month s, j indicates the start of the next scheduled FOMC meeting, and $j+$ indicates the close of the meeting. It follows that θ_{j+} is the value of the Fed target at the close of the meeting. To extract a shock we write eq(4.31)

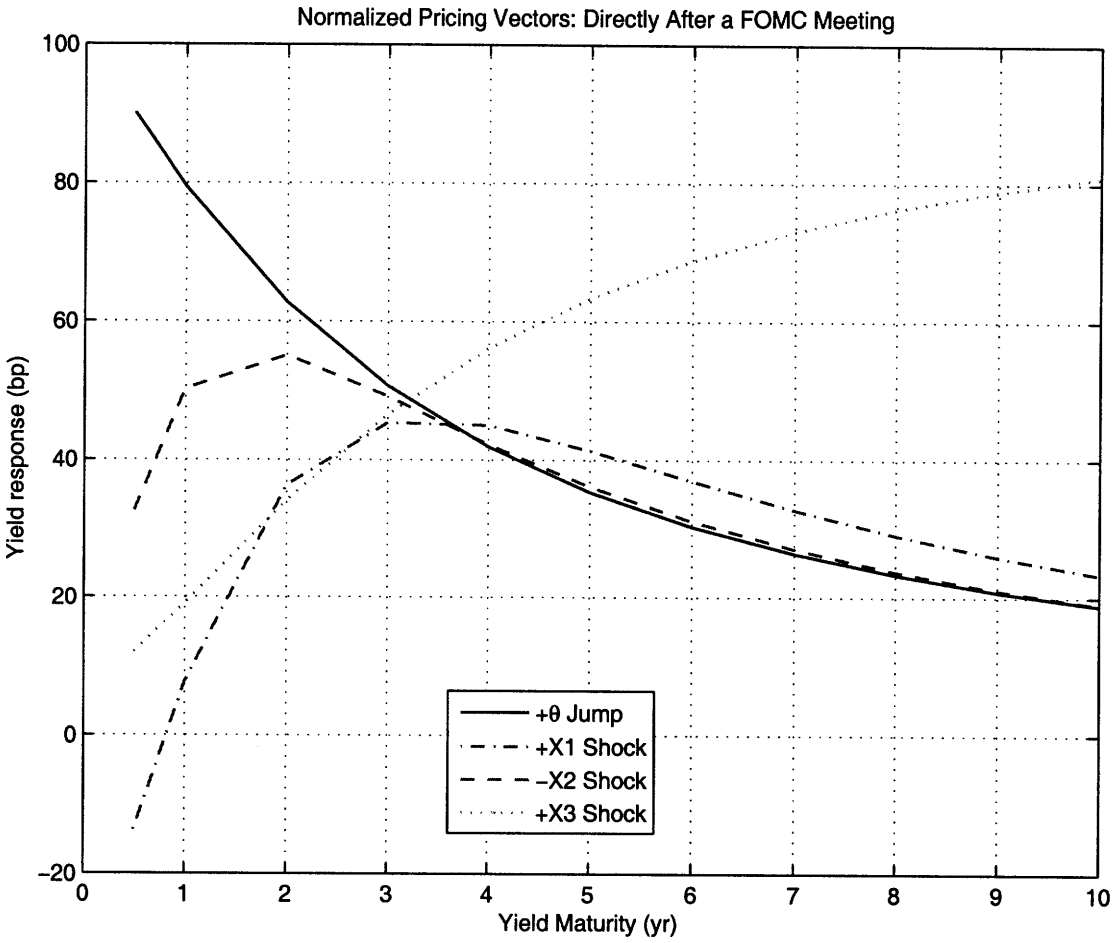

Figure 4-18: Normalized pricing vector. C_X is normalized to display the response to a 1 standard deviation shock. C_θ requires no such normalization.

under P with an explicit risk premium.

$$f_t^s = \frac{j}{m}\theta_t + \frac{m-j}{m}E_t^P[\theta_{j+}] + \mu_t^s \tag{4.32}$$

$$= \theta_t + \frac{m-j}{m}E_t^P[\Delta\theta_{j+}] + \mu_t^s \tag{4.33}$$

where μ_t^s is the time t risk premium for the month s contract, and $\Delta\theta_{j+}$ is the target change which occurred during the meeting. The expected change can now be written as

$$E_t^P[\Delta\theta_{j+}] = \frac{m}{m-j}(f_t^s - \theta_t) - \frac{m}{m-j}\mu_t^s \tag{4.34}$$

The first term is the anticipated change we seek to measure, the second term is the risk premium which corrupts the signal. The risk premium corruption is magnified as $j \rightarrow m$. For completeness we give the associated monetary policy shock as

$$\Delta\theta^u_{j+} = \frac{m}{m-j}(f^s_t - \theta_t) - \frac{m}{m-j}\mu^s_t - \Delta\theta_{j+} \tag{4.35}$$

where $\Delta\theta^u_{j+}$ denotes the unanticipated change in θ. An alternative method which attempts to remove the risk premium of eq(4.35) is to measure the shock component directly as first presented in [45].

If a change in the Fed target is fully anticipated, then the price of the associated Fed future contract over the same period should remain constant. Assume the market fully anticipates a 25 basis point increase at the next FOMC meeting which is scheduled to occur tomorrow. Since the Fed future contracts are based on the future expected average target rate, the 25 bp increase should be already priced in. It follows that any change in a Fed future contract across scheduled meetings represents a revision in the expected target level for the remainder of the contract month. This revision is a scaled measure of the monetary policy shock. Specifically we can define monetary policy shocks as

$$\Delta\theta^u_{j+} = \frac{m}{m-j}(f^s_{j+} - f^s_j) \tag{4.36}$$

where f^s_j is the Fed future contract immediately before a scheduled meeting, f^s_{j+} is the contract value immediately after the meeting, and s is the current month contract. The critical assumption is the risk premium at time j is approximately equal to the risk premium at time $j+$. Clearly a constant risk premium would satisfy this assumption. Even if the risk premium is time varying, the variation should be small over the course of one day.

Figure 4-19 plots monetary policy shocks derived from the model, as well as shocks constructed via eq(4.36). We stress that the model calibration made no use of information in Fed future contracts. The two measures show an impressive overlap in information content as measured by a 0.7632 sample correlation coefficient. The two

time series differ the most during the tightening cycle which began June 30th 2004. During this period market participates widely expected the FOMC to raise the target by 25 basis at each consecutive meeting. This constant jump expectation poses challenges to model dynamics.

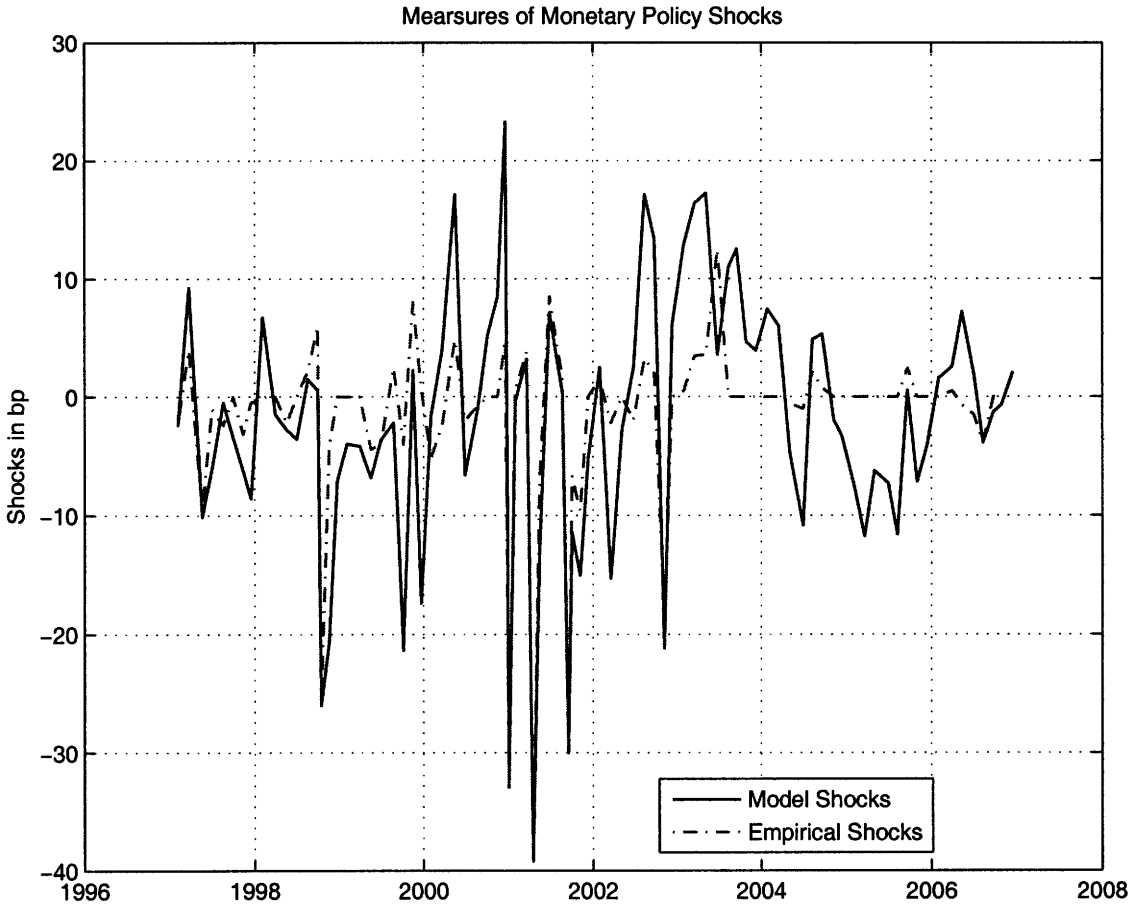

Figure 4-19: Time series of model implied monetary policy shocks, as well as shocks measured via Fed future contracts. Sample correlation of two measures is 0.7632.

Given measures of monetary policy shocks, we turn to the task of measuring the effect of shocks on yields. We construct an empirical estimate of yield response to policy shocks via a linear regression, and compare the results to the yield response imbedded in our constructed model. We find the results to be almost identical.

To construct the effect of shocks on yields within the model framework, we look

to the model pricing formula.

$$Y(t,T) = C_o(t,T) + C_\theta(t,T)\theta_t + C_X(t,T)X_t + \epsilon_t \tag{4.37}$$

Changes in yields are expressed as

$$\begin{aligned} Y(t+1,T+1) - Y(t,T) &= [C_o(t+1,T+1) - C_o(t,T)] \\ &+ [C_\theta(t+1,T+1)\theta_{t+1} + C_\theta(t,T)\theta_t] \\ &+ [C_X(t+1,T+1)X_{t+1} - C_X(t,T)X_t] \\ &= [C_o(t+1,T+1) - C_o(t,T)] \\ &+ [C_\theta(t+1,T+1)E_t[\theta_{t+1}] + C_\theta(t,T)\theta_t] \\ &+ [C_X(t+1,T+1)X_{t+1} - C_X(t,T)X_t] \\ &+ C_\theta(t+1,T+1)\Delta\theta^u_{t+1} \end{aligned} \tag{4.38}$$

where $\Delta\theta^u_{t+1}$ is the unanticipated change to the target at $t+1$. It follows that $C_\theta(t+1,T+1)$ defines how model yields respond to shocks.

For an empirical reference we follow [45] and fit weekly yields to a linear equation in the expected and unexpected changes in the target rate.

$$\Delta Y_t = \alpha + \beta_e \Delta\theta^e_t + \beta_u \Delta\theta^u_t + \epsilon_t \tag{4.39}$$

where $\Delta\theta^e_t$ is the expected target change and $\Delta\theta^u_t$ is the unexpected target change. Most studies treat jump events separate from non-jump events, estimating eq(4.39) from weekly samples where the target actually changed. The intuition typically being that markets respond differently to these distinct events. We estimate the parameters of eq(4.39) via ordinary least squares (OLS). Table 4.7 reports the parameter estimates using weekly data. The results are consistent with the results given in [45].

Figure 4-20 compares the model's ability to transmit monetary policy shocks to yields. We find the model shows very close agreement with our empirical estimates.

Maturity (yr)	α	β_e	β_u	R^2
0.5	0.53	11.62	98.76	0.75
	(0.8)	(3.2)	(7.6)	
1	0.9	9.74	78.46	0.49
	(0.9)	(1.9)	(12)	
2	0.89	6.33	58.74	0.23
	(0.7)	(1)	(6.9)	
3	0.66	3.76	51.38	0.17
	(0.5)	(0.6)	(5.2)	
4	0.41	2.73	45.98	0.13
	(0.3)	(0.4)	(4.2)	
5	0.05	2.17	42.08	0.11
	(0.01)	(0.3)	(3.3)	
6	-0.13	1.71	37.79	0.09
	(0.1)	(0.3)	(2.9)	
7	-0.33	0.88	32.02	0.06
	(0.2)	(0.1)	(2.2)	
8	-0.57	1.09	29.39	0.06
	(0.4)	(0.2)	(2)	
9	-0.69	0.48	25.64	0.04
	(0.5)	(0.1)	(1.7)	
10	-0.63	1.01	24.11	0.04
	(0.4)	(0.1)	(1.6)	

Table 4.7: OLS estimates of the coefficients in eq(4.39). Heteroskedasticity-consistent White t-statistics in parentheses.

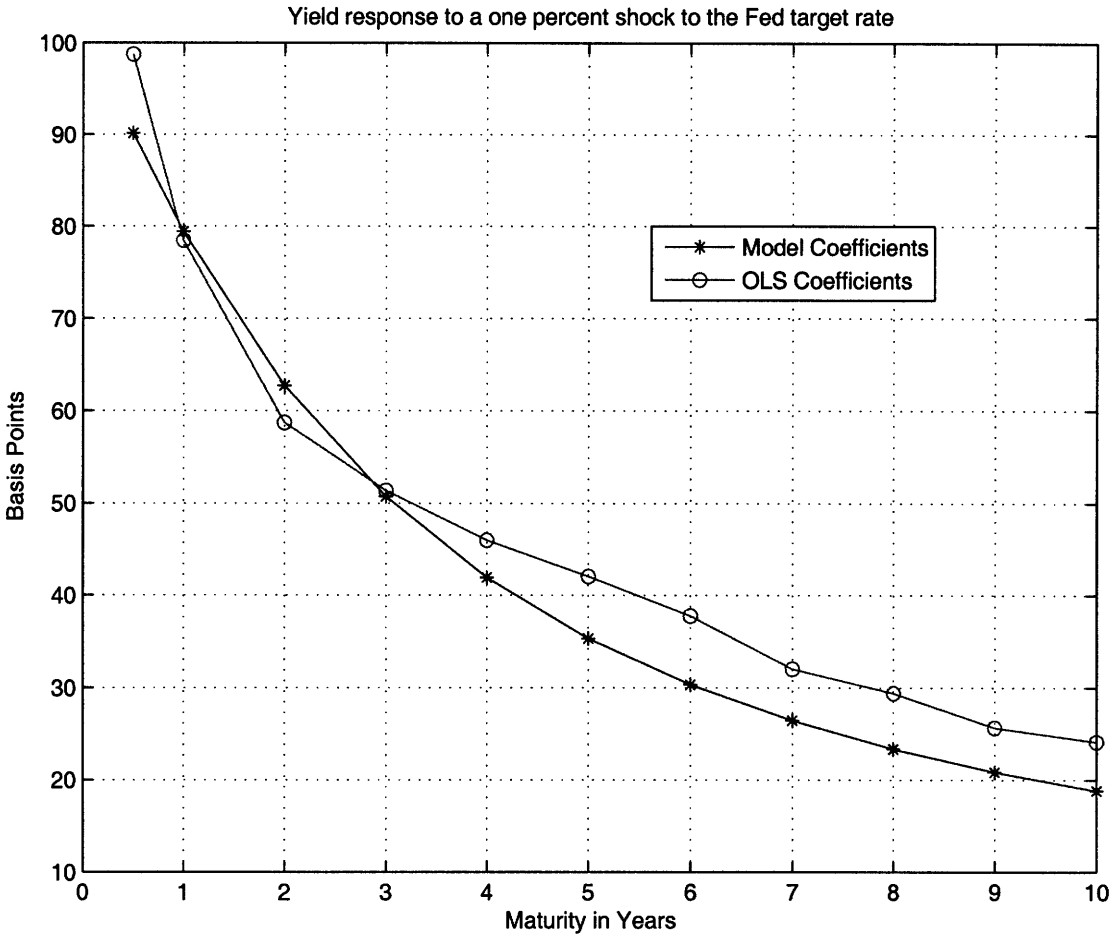

Figure 4-20: Response of Yields to shocks in monetary policy, as measured by unanticipated changes in the Fed target rate.

4.4 Risk Premium

A primary driver to conduct term structure research within a Dynamic Term Structure Model (DTSM) framework is the ability to easily identify risk premium. In this section we focus on the model's ability to explain one year excess bond returns. We find the inclusion of the Fed target as a state variable greatly improves the model's predication of excess return. We find evidence that the improvement is due to the target as a proxy for short dated rates. We also find substantial improvement when the information embodied in the term structure of target rates is incorporated during the estimation process. Thus the target also serves as a conduit to integrate market participates expectations for future target levels. Linear regressions provide support to the risk premium view of the target as a proxy for a short dated rate and a conduit for the term structure of target rates.

Background

Define one year excess return on a n year bond as

$$xr(t+1,n) = nY(t,n) - (n-1)Y(t+1,n-1) - Y(t,1) \tag{4.40}$$

where $Y(t,n)$ is the time t yield on an n year bond. Figure 4.4 plots the realized one year excess return for our data set, using overlapping windows. In summary the one year excess return is the return on a bond position above the one year yield available at time t. The expectation hypothesis states that forward rates are rational expectations of future yields, and as such no time t information should predict excess bond returns. In recent years several empirical studies have presented evidence rejecting the expectation hypothesis by finding evidence of predictability in excess bond return.

There exists a rich set of literature which explores the predictability of bond returns. One of the first studies [5] found the spread between forward rates and one year yields predicted excess bond returns. [7] later showed the slope of current yield increased this predictability. More recently a study in [14] showed that including all

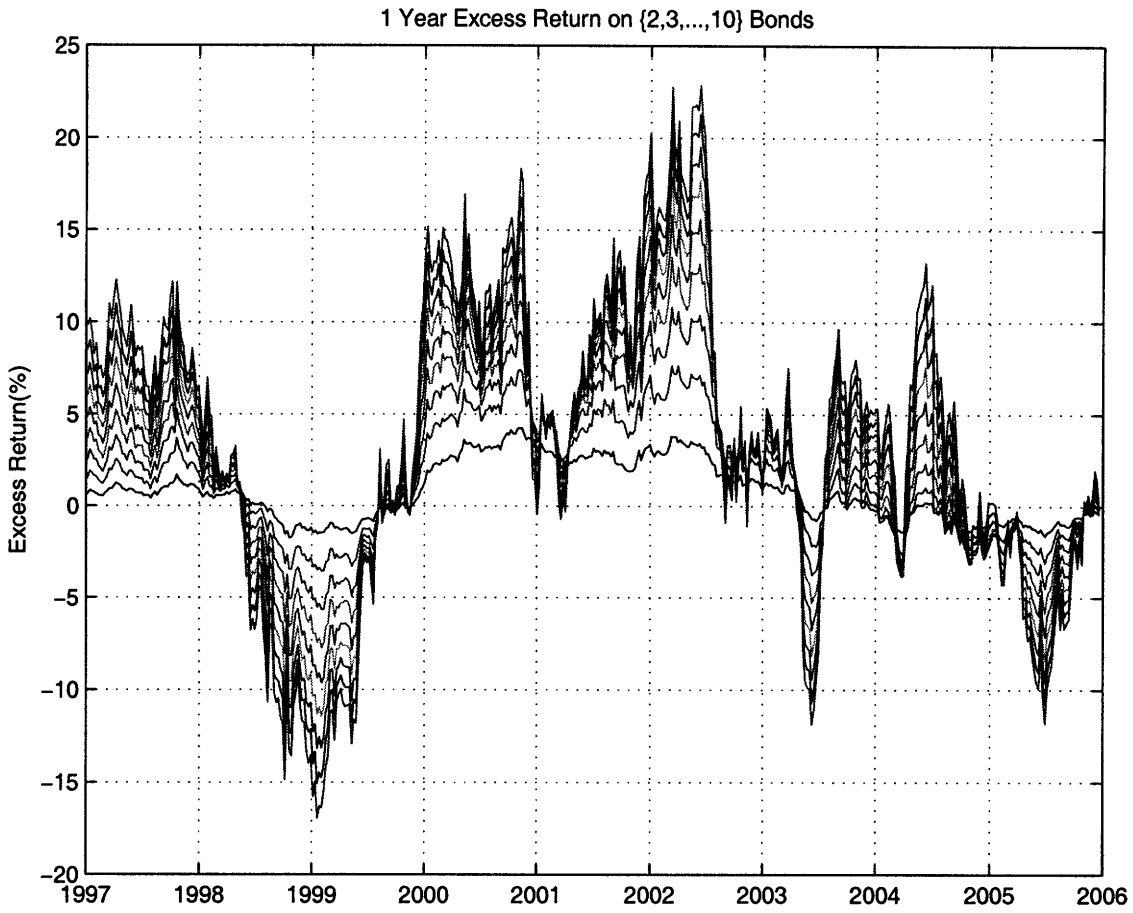

Figure 4-21: Realized one year excess return on $\{2, 3, ..., 9, 10\}$ year bonds using weekly sampled bond yields. Return calculations use overlapping windows.

time t information in forward rates is important in explaining bond returns.[2].

We seek a model expression for the expected value of the return, or $E_t[xr(t+1, n)]$. Where we stress the expectation is under the historical P measure. Within a DTSM framework we can express the three terms of eq(4.40) as

$$nY(t,n) = -\log E_t^Q \left[\exp\left(-\int_t^{t+n} r(\tilde{X}_s) ds \right) \right] \quad (4.41)$$

$$-(n-1)E_t^P[Y(t+1,n-1)] = E_t^P \left[\log E_{t+1}^Q \left[\exp\left(-\int_{t+1}^{t+n} r(\tilde{X}_s) ds \right) \right] \right] \quad (4.42)$$

$$-Y(t,1) = \log E_t^Q \left[\exp\left(-\int_t^{t+1} r(\tilde{X}_s) ds \right) \right] \quad (4.43)$$

[2][14] extend this to show a single factor, which is affine in yields, explains the majority of excess bond return.

Combine eq(4.41) and eq(4.43) to find

$$nY(t,n) - Y(t,1) = -\log E_t^Q \left[\exp\left(-\int_{t+1}^{t+n} r(\tilde{X}_s) ds \right) \right] \quad (4.44)$$

$$\approx C_{\tilde{X}}(t+1, n-1) E_t^Q(\tilde{X}_{t+1}) \quad (4.45)$$

where we ignore terms related to the Jensen inequality. We remind readers that though our state space is not affine, we presented a method to construct affine pricing coefficients in section 3.2.4. A similar approximation for eq(4.42) results in

$$-(n-1) E_t^P [Y(t+1, n-1)] \approx C_{\tilde{X}}(t+1, n-1) E_t^P(\tilde{X}_{t+1}) \quad (4.46)$$

It follows that

$$E_t^P [xr(t+1,n)] \approx C_{\tilde{X}}(t+1, n-1) \left[E_t^P(\tilde{X}_{t+1}) - E_t^Q(\tilde{X}_{t+1}) \right] \quad (4.47)$$

Eq(4.47) highlights how risk premium is specified and transmitted to bond returns. The differences in drifts under each measure specifies the associated risk premium over any period of time. The pricing coefficients $C_{\tilde{X}}$ provide the weighting of each bond to the risk premium. An alternative to using the approximation in eq(4.47) is to compute the expected value of $Y(t+1, n-1)$ via Monte Carlo and evaluate eq(4.40) directly.

Model Performance

Using the model parameters of table 4.1 we evaluate eq(4.47) for overlapping one year expected excess bond returns. Monte Carlo simulations confirm the errors induced by ignoring the Jensen terms are well contained, and do not change the ensuing results. The resulting model expected excess return is plotted in figure 4.4. We find the model risk premium displays an ability to switch sign, due to the flexible specification of section 3.2.3. Comparing the model expected excess return to the realized excess return we find good tracking. With the exception of a window around 1999, the model matches the level of risk premium quite well. In particular the model reflects

the negative dips in 2003 and 2005.

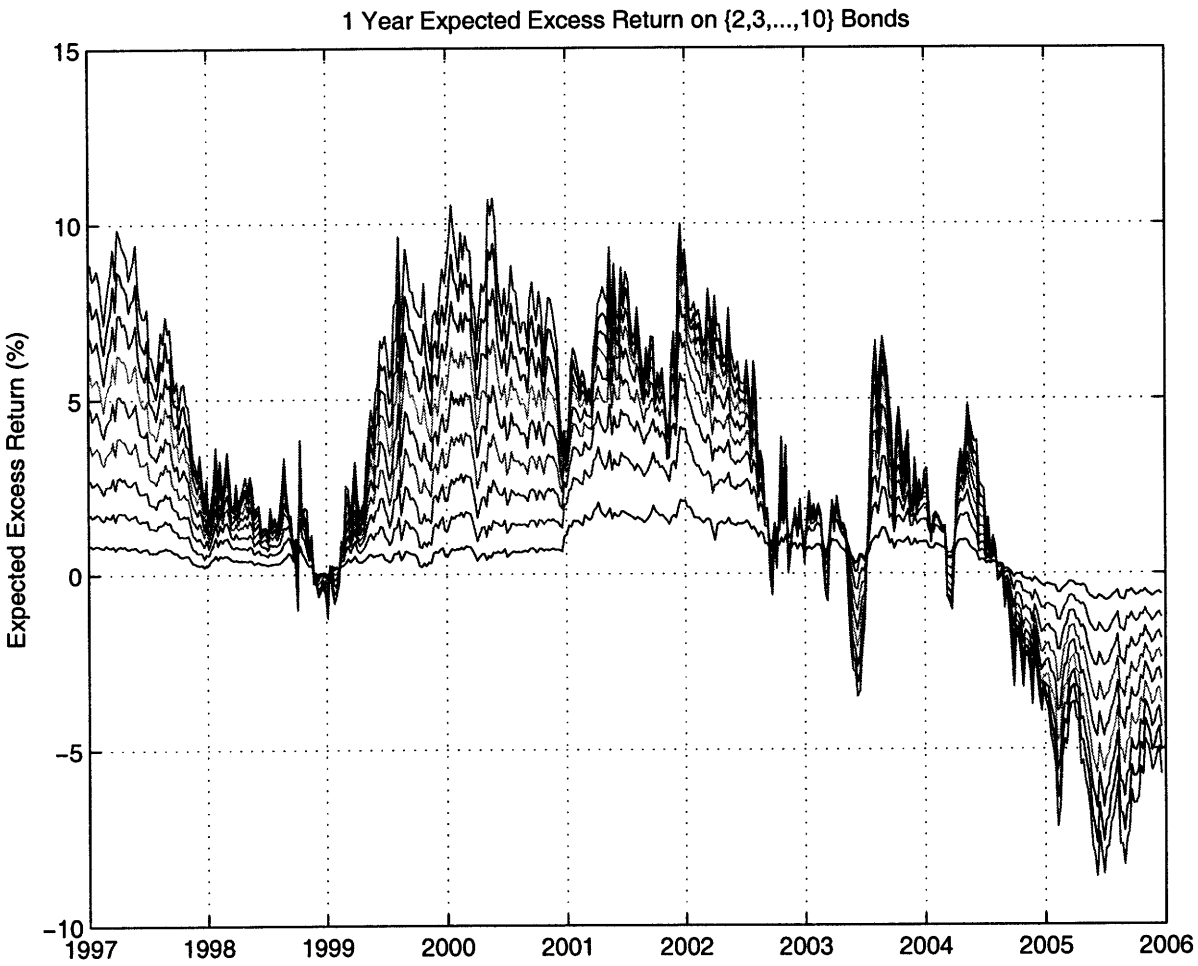

Figure 4-22: Model expected one year excess return on $\{2, 3, ..., 9, 10\}$ year bonds using weekly sampled bond yields. Return calculations use overlapping windows.

To quantify the model's ability to reflect risk premium we use the coefficient of determination or R^2. To better view the ensuing result we require a benchmark. Since the purpose of our study is to investigate the effect of including the target in an ATSM model we choose the $A_1(3)$ ATSM as a benchmark. As detailed in section 3.2.1 this is essentially our model without the target as a state variable. We calibrate the model parameters using the same estimation method and weekly data set. A full description of the model with estimated model parameters is available in Appendix C.1.

Figure 4.4 plots the results. *Ylds Only* indicates the we are using the model pa-

rameters of table 4.1, which were estimated using the yield data. We note a significant increase in performance with respect to the benchmark. We find a significant increase in explanatory power up to the six year node, peaking at an additive five percent in explained variation. The increase in performance begins to decline at the six year node, converging to zero at ten years. The performance is consistent with many of our previous findings and intuition. Figure 4-20 showed the response of yields to shocks is strongest for shorted dated maturities. Figure 4-13 implies the strongest link to monetary policy comes from the one and two year nodes. Section 4.3 detailed that the dominating component of target jump intensity X_2 displays a curvature effect, peaking at the two year node. These results are also in line with the discussion of the target and it's role in adjusting the monetary base of section 1.

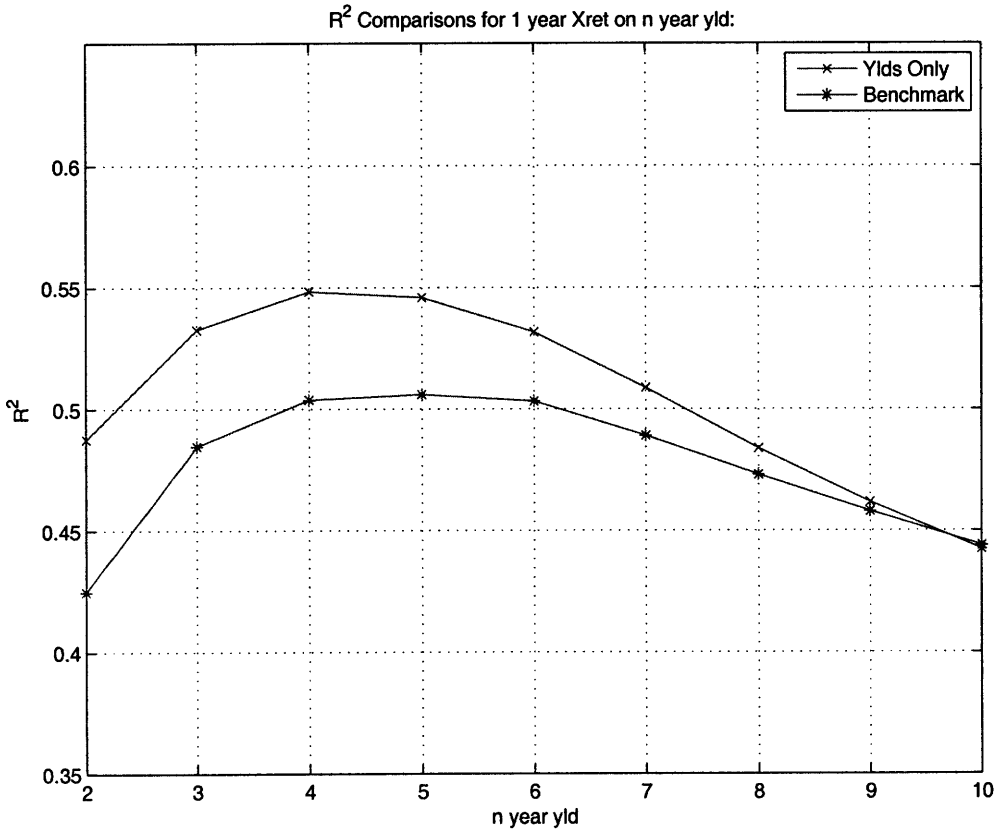

Figure 4-23: Explaining excess bond returns. The Benchmark is a $A_1(3)$ model with no jump component. Ylds indicates the use of model parameters of table 4.1, which are estimated using yields only.

Term Structure of Target Rates

We are also interested in determining if information contained in the Fed future contracts contain any risk premium information not already in yields. Recall section 4.2 presented small out-of-pricing errors for the rolling time series of Fed future contracts. When the time series is incorporated during estimation the resulting pricing errors are essentially unchanged, as are the other performance metrics discussed in section 4.1 and section 4.3. The associated R^2 with respect to excess bond returns is also relatively unchanged. Though there does seem to be a slight shifting of explanatory power to the lower end of the yield curve.

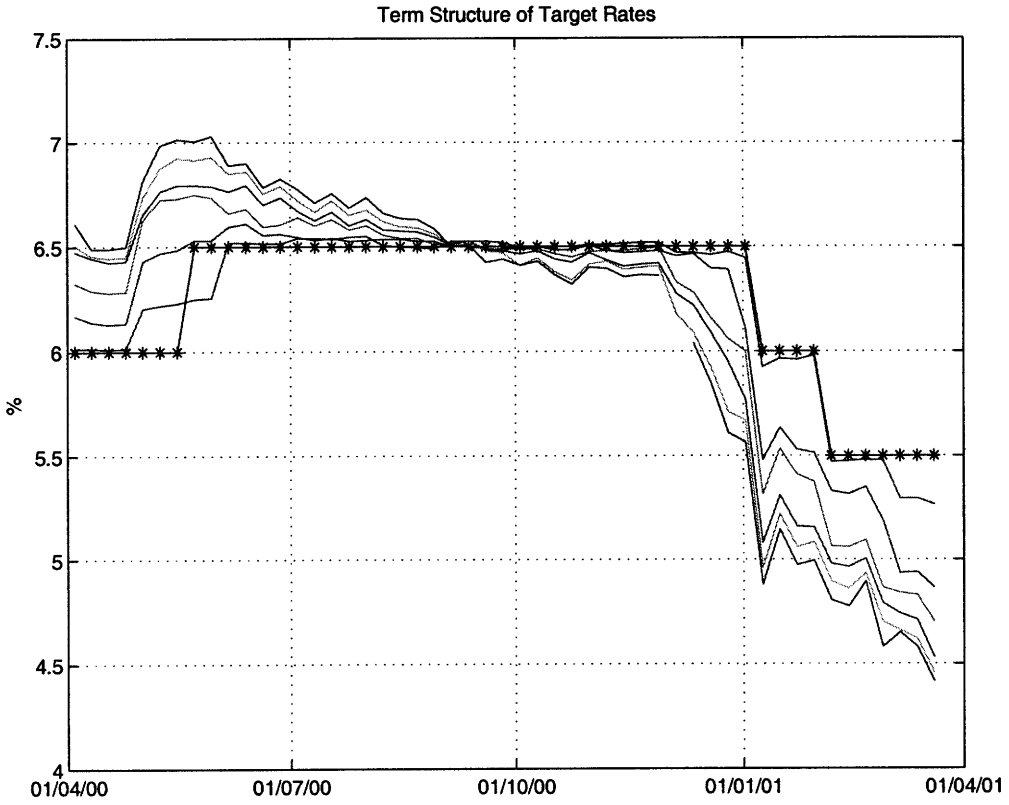

Figure 4-24: Plot of the term structure of target rates taken from Fed future contracts during the 2000 turning point. Doted line is the actual Fed target rate.

We now turn to the entire term structure of target rates. Define the term structure of target rates to be the expected level of the target rate over the next the next twelve months. This information is easily extracted from the currently traded Fed future

contracts. Figure 4.4 and 4.4 plot the full span of future contracts for two distinct time periods. Each graphically demonstrates the additional information contact in the additional contracts. We see in figure 4.4 that while the near dated contract is essentially flat, the far dated contracts embed the market's believe of a turning point in the target rate. In figure 4.4 we find the market belief that the Fed will tighten in consistent but prolonged steps. It follows that the market's short term prediction of the future target level may help predict excess return.

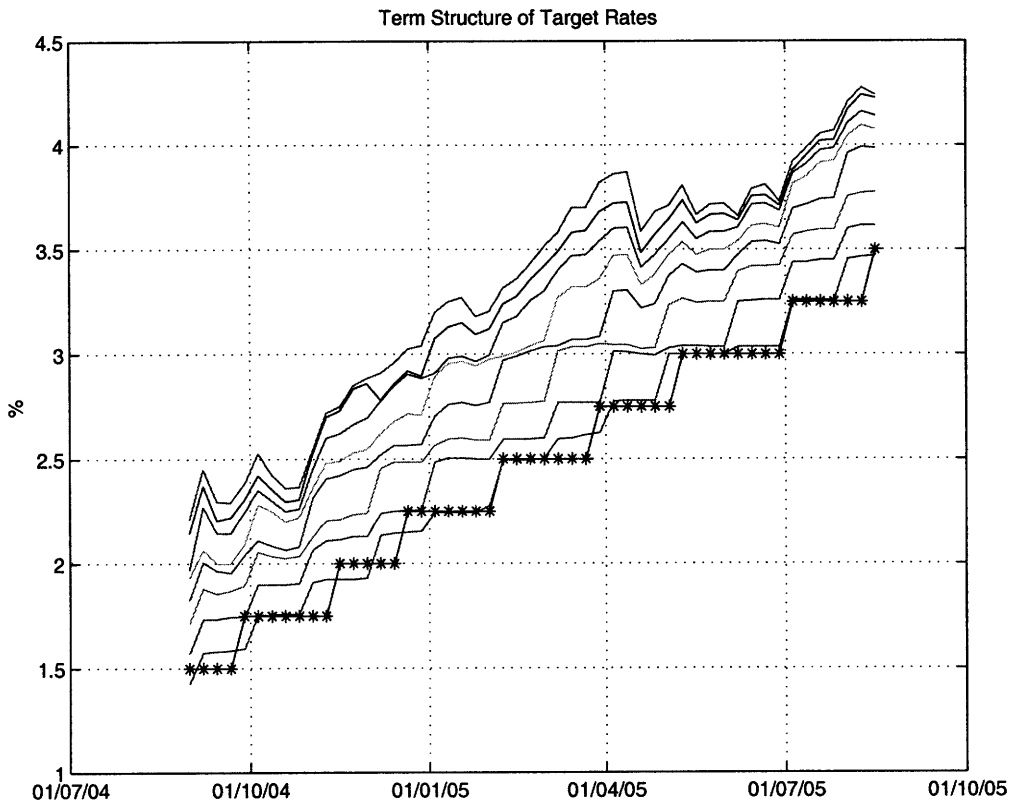

Figure 4-25: Plot of the term structure of target rates taken from Fed future contracts during the 2005 tightening cylce. Doted line is the actual Fed target rate.

We aim to include this information content by incorporating the long dated contracts during the estimation process. Recall from section 4.2 that our model price for a Fed future contract is the daily average value of the target rate during the contract month. Fed future contracts trade for twelve months out, where open interest is often thin for later months. To avoid spurious contract prices, we remove any contract

who's open interest is less then ten percent of the total open interest. This normalization ensures compensation for the six fold increase in total open interest since 2000. Assume each contract possesses additive error identical to eq(4.11). Then the model pricing equation for a Fed future contract for month s is

$$f_t^s = \frac{1}{m}\sum_{i=1}^{m} E_t[\theta_i] + \epsilon_t \tag{4.48}$$

where there are m days in month s. These contracts are priced in the model via the techniques described in section 4.2.

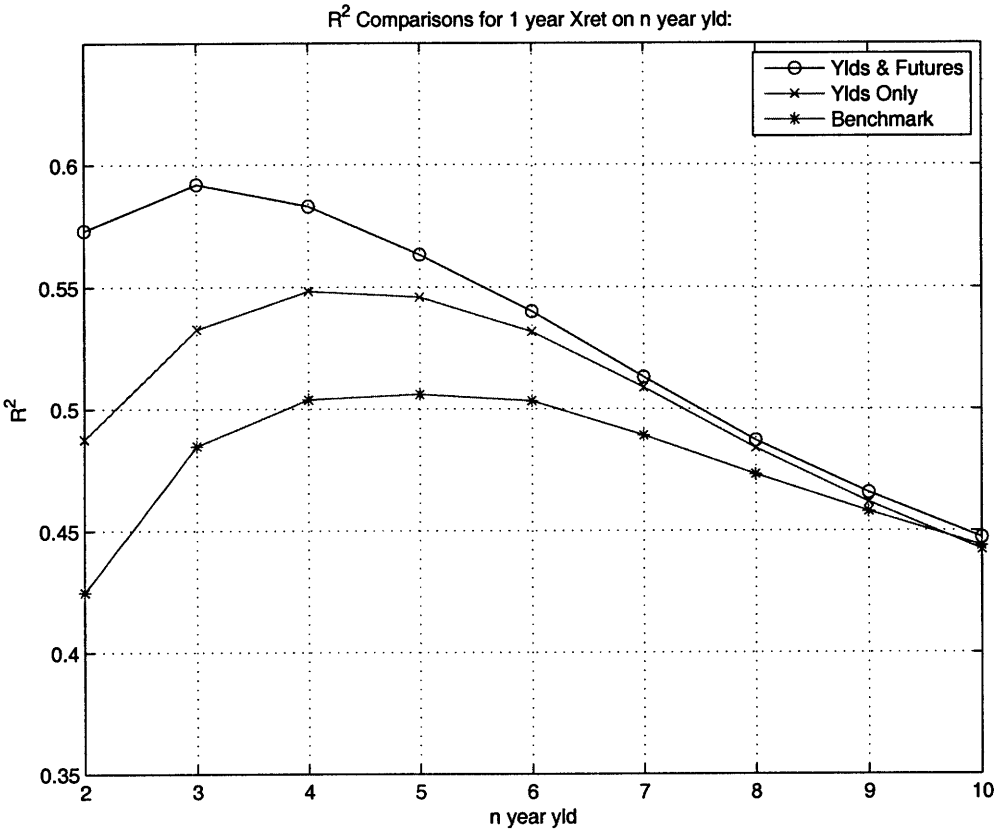

Figure 4-26: Explaining excess bond returns. The Benchmark is a $A_1(3)$ model with no jump component. Ylds indicates the use of model parameters of table 4.1, which are estimated using yields only. Ylds & Futures represents model performance when yields and Fed future contracts are used to calibrate model parameters.

The resulting parameter estimates and in-sample pricing errors are in appendix C.2. Comparing the errors in table C.3 to the errors of table 4.6, we conclude there is

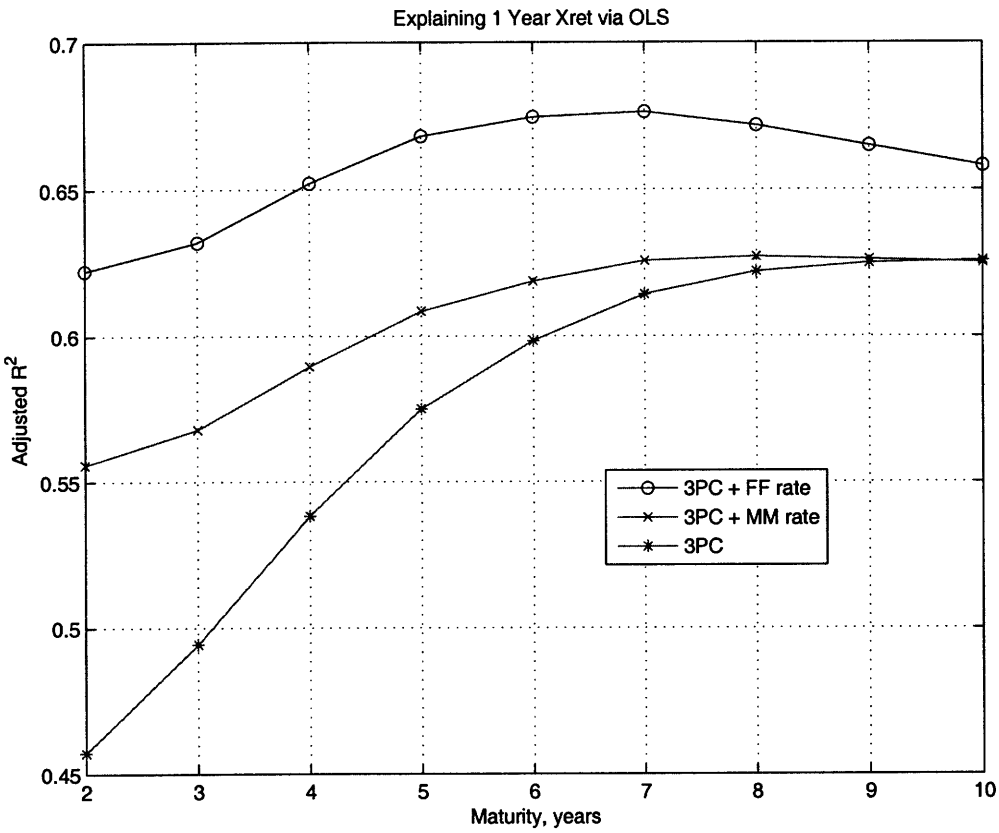

Figure 4-27: Explaining excess bond returns via ordinary lease squares regression. 3PC are the three principle components of all yields of section 3.4.2. MM indicates the short dated 1 week money market rate. FF indicates the 1 month ahead Fed Future contract as described in section 3.4.3.

no significant improvement to pricing errors when including the Fed future contracts. However when we use the estimates of table C.2 to compute risk premium we find substantial improvement. Figure 4-26 plots the associated R^2 metrics for the three cases we've discussed. For clarity we stress *Ylds & Futures* represents model performance when yields and Fed future contracts are used to calibrate model parameters. We find substantial improvement in the model's ability to predict excess bond return. This result implies there is information in the term structure of Fed future contracts which while it does not assist in pricing bonds, does help in predicting returns on bonds. We stress that the future expected target changes imbedded in Fed futures is under the pricing measure. We find the contracts especially at long dated intervals display strong risk premium, which is supported by empirical studies [53].

Based on the above analysis we postulate that the improved predictability is due to the target as a proxy for a short dated yield, and as a conduit for future expected target changes. In an effort to support this view we turn to more empirical estimates via linear regressions. Specifically we construct the first three principle components of weekly yields, the one week money market rate, and a mid-dated Fed future contract. We regress the realized excess return on this data and plot the associated R^2 in figure 4.4. Comparing figure 4.4 to 4-26 we find many similarities. The relative spreads at the short to mid maturities nodes displays excellent tracking. Unfortunately the long dated maturities do not follow this trend. In summary figure 4.4 provides support to the view that the improved predictability is due to the target as a proxy for a short dated yield, and as a conduit for future expected target changes.

Chapter 5

Conclusion

This thesis investigated the implications of explicitly modeling the monetary policy of the Central Bank within a Dynamic Term Structure Model (DTSM). We followed Piazzesi (2005) and implemented monetary policy by including the Fed target rate as a state variable. A non-linear switching process, was found to accurately model the target dynamics while allowing for restrictions to ensure tractability. The flexible risk specification of Cheridito et al (2007) was incorporated, with extensions to the target jump process explored. Model parameters were estimated via a simulated maximum likelihood estimation scheme with importance sampling. Finally a Bayesian particle filter was found to be a useful robustness check.

Our results support those in Piazzesi (2005), revealing a substantial improvement in pricing errors especially on the short end of the yield curve. The model construction was shown to provide a natural framework to inspect monetary policy information embedded in yields, which was found to be substantial. We found the addition of the target rate improves the model's ability to explain excess return. An ability which is increased with the inclusion of the full term structure of target rates, as measured from Fed future contracts. We presented a view that the improved performance is due to the target as a proxy for short term rates, and a conduit to express the information content of the term structure of target rates.

Appendix A

Bond Pricing Accuracy Check

Two approximations were presented in section 3.2.4. The first concerned the non-linear terms in the Feynman-Kac PDE of eq(3.22) which prevent the construction of the ODEs, the essence of tractability in our problem. In section 3.2.4 we applied the following Taylor Series expansion to achieve tractability.

$$\Big|\lambda(X)\Big|\Big[B(t,T,X,\theta+\text{sign}(\lambda(X))J_\theta)-B(t,T,\widetilde{X})\Big]\approx J_\theta\lambda(\widetilde{X})C_\theta(t,T)B(t,T,\widetilde{X}) \tag{A.1}$$

The second concerned the normalization of scheduled FOMC meetings, by using the true meeting grid for the first FOMC meeting and uniform spacing for all which follow. The primary motivation for this approximation is computational efficiency. However the meeting schedule for maturities over two years is never known to market participates when prices are set. As such a normalization appears in line with market realities.

We test both approximations simultaneously via a Monte Carlo estimation of bond prices. We use the values of inferred state variables as initial conditions, and take one draw by simulating the state space for ten years. Bond prices are then simple Monte Carlo expectations. Table A.1 displays the results. MC-ODE is the RMSE of the difference between the MC yields and the ODE yields.

Maturity (yr)	Monte Carlo RMSE (bp)	ODE RMSE (bp)	MC-ODE (bp)
0.5	1.2076	0	1.2076
1	4.4001	4.3132	1.1404
2	1.2896	0	1.2896
3	2.0585	1.2804	1.4168
4	3.641	2.8824	1.4644
5	4.4284	3.6166	1.5333
6	4.6089	3.6875	1.6653
7	4.1519	3.0364	1.7637
8	3.4201	2.2263	1.8526
9	2.5737	0.99687	1.9727
10	2.0663	0	2.0663

Table A.1: Monte Carlo estimates of pricing errors due to linearization of the jump term and normalization of the meeting schedule. Root mean squared errors (RMSE) in basis points (pb). MC-ODE is the RMSE of the difference between the MC yields and the ODE yields.

Appendix B

Estimation

B.1 Simulated Maximum Likelihood

Simulated Maximum Likelihood [49] is a popular method for estimation of continuous time processes from discretely sampled data. Consider a continuous time drift diffusion process $X(t)$, who's dynamics are specified via the following stochastic differential equation (SDE)

$$dX(t) = a(X(t), \gamma)dt + b(X(t), \gamma)dW(t) \tag{B.1}$$

where W is a standard Brownian Motion defined in a complete probability space, a is a specified drift function, b is a specified diffusion function, and γ ia a parameter vector. Conditions on a, b, and γ to ensure the existence and uniqueness are stated in [49], and discussed at detail in [48].

Assume X is sampled at $N+1$ discrete points denoted as $X_{(N)} = (X_o, X_1, X_2, \ldots, X_N)$. Assume a joint density $f_X(X_{(N)}; \gamma)$, with corresponding continuously differentiable likelihood function $L_X(X_{(N)}; \gamma)$. Since X is a Markov process we write the likelihood function as

$$L_X(X_{(N)}; \gamma) = f(X_o; \gamma) \prod f_X(X_{n+1}, t_{n+1} | X_n, t_n, \gamma) \tag{B.2}$$

In general the transition densities in the product are not known, and must be approximated. The SML method approximates the transition densities by simulating stochastic paths between observations. Take the the transition density $f_X(X_{n+1}, t_{n+1} | X_n, t_n, \gamma)$

as an example. Discretize time between t_n and t_{n+1} into M subintervals of length h. The transition density between these subintervals is then approximated as Gaussian. The transition density for the first subinterval is

$$f_X(x, t_{n+h}|X_n, t_n, \gamma) \approx \phi(x; X_n + ha(X_n, t_n), hb(x, t_n)^2) \tag{B.3}$$

where $\phi(\cdot; M, V)$ is the Gaussian density with mean M and variance V. The approximated transition density is then applied to subintervals up to the $M^{th} - 1$ subinterval. At which point the approximation becomes

$$f_X(X_{n+1}, t_{n+1}|\widehat{X}, t_{n+1}-h) \approx \phi(X_{n+1}; \widehat{X}+ha(\widehat{X}, t_{n+1}-h), hb(\widehat{X}-h, t_{n+1}-h)^2) \tag{B.4}$$

Where $\widehat{X}$ is the simulated value at the indicated time. The estimate for the transition density from t_n to t_{n+1} is then taken via an average of eq(B.4). That is for J simulated sample paths we find

$$\hat{f}_X(X_{n+1}, t_{n+1}|X_n, t_n) = \frac{1}{J}\sum_{j=1}^{J} \phi(X_{n+1}; \widehat{X} + ha(\widehat{X}, t_{n+1} - h), hb(\widehat{X}, t_{n+1} - h)^2) \tag{B.5}$$

The final expression in eq(B.5), can also be viewed as constructing the expectation of a random variable $z = f_X(X_{n+1}, t_{n+1}|z, t_{n+1}-h)$, with a density drawn from simulated paths and equal to $f_Z(z, t_{n+1}-1|X_n, t_n)$. Details on convergence of the SML estimator to the true likelihood estimator is given in [49].

Appendix C

Details on Model Extensions

C.1 $A_1(3)$ Benchmark

In the language of [19] the benchmark model of section 4.4 is an $A_1(3)$ model. This is essentially the continuous component of the state space presented in section 3.2.1. We leverage the same estimation scheme presented in 3.3, using the weekly data set of section 3.4. We summarize here the model construction and estimation results.

The dynamics under P are given as

$$dX_t = \mu^P(X_t)dt + \sigma(X_t)dW_t^P \tag{C.1}$$

where

$$\mu^P(X_t) = K_o^P + K_1^P \cdot X_t = \begin{bmatrix} k_{01}^P \\ k_{02}^P \\ k_{03}^P \end{bmatrix} + \begin{bmatrix} k_{11}^P & 0 & 0 \\ k_{21}^P & k_{22}^P & k_{23}^P \\ k_{31}^P & k_{32}^P & k_{33}^P \end{bmatrix} \cdot X_t \tag{C.2}$$

and

$$\sigma(X_t) = \begin{bmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1 + b_{21}X_t^1} & 0 \\ 0 & 0 & \sqrt{1 + b_{31}X_t^1} \end{bmatrix} \tag{C.3}$$

We employ the same flexible risk specification of section 3.2.3. This provides the parametric form of dX under Q as

$$dX_t = \mu^Q(X_t)dt + \sigma(X_t)dW_t^Q \tag{C.4}$$

where

$$\mu^Q(X_t) = K_o^Q + K_1^Q \cdot X_t = \begin{bmatrix} k_{01}^Q \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} k_{11}^Q & 0 & 0 \\ k_{21}^Q & k_{22}^Q & k_{23}^Q \\ k_{31}^Q & k_{32}^Q & k_{33}^Q \end{bmatrix} \cdot X_t \tag{C.5}$$

We define the short rate as an affine function of the state space

$$r(X_t) = \rho_o + \rho_1 X_1 + \rho_2 X_2 + \rho_3 X_3 \tag{C.6}$$

For admissibility we require

1. The two zeros in the first row of K_1^P
2. $k_{01}^P, k_{01}^Q \geq 0.5$
3. $b_{j1} \geq 0$ for $j = 1, 2$

For identification we require

1. The two zeros in K_o^Q
2. $\rho_j \geq 0$ for $j = 2, 3$
3. $k_{22}^Q > k_{33Q}$

To construct the $A_1(3)$ pricing equations we substitute the above dynamics into

the PDE of eq(2.6). The resulting system of ODEs is

$$\frac{dC_o}{dt} = \rho_o - C_{X1}K_{01} - \frac{1}{2}\left(C_{X2}^2 + C_{X3}^2\right) \tag{C.7}$$

$$\frac{dC_{X1}}{dt} = \rho_1 + C_{X1}K_{11} + C_{X2}K_{21} + C_{X3}K_{31} - \frac{1}{2}\left(C_{X1}^2 + b_{21}C_{X2}^2 + b_{31}C_{X3}^2\right) \tag{C.8}$$

$$\frac{dC_{X2}}{dt} = \rho_2 + C_{X2}K_{22} + C_{X3}K_{32} \tag{C.9}$$

$$\frac{dC_{X3}}{dt} = \rho_3 + C_{X2}K_{23} + C_{X3}K_{33} \tag{C.10}$$

Since we no longer track the FOMC meeting schedule, the pricing equations are time invariant, that is $C(t,T) \rightarrow C(n)$ for $n = T - t$. Table C.1 displays the point estimates of the static model parameters. Unfortunately the small sample leads to the vast majority of estimates being insignificant from zero at the 5% level.

C.2 Term Structure of Target Rates

Section 4.4 revealed that information embedded in the full term structure of target rates improves the model's predictability of excess return. Where the term structure of target rates are implied by the presently traded Fed future contracts. To incorporate this information into the model, we assume each of the

Table C.2 contains the estimates of all model parameters, when all available Fed future contracts priced. Table C.3 contains the associated in-sample pricing errors on yields.

Parameter	Estimate	Standard Error	t-Statistic
k_{o1}^{Q}	2.6599	0.9772	**2.7219**
k_{11}^{Q}	0.6013	0.0070	**85.7945**
k_{21}^{Q}	698.8810	4840.9453	0.1444
k_{22}^{Q}	1.3683	0.0263	**52.0106**
k_{23}^{Q}	-2.3808	7.7402	-0.3076
k_{31}^{Q}	-1.4796	12.6230	-0.1172
k_{32}^{Q}	0.0183	0.0594	0.3089
$k_{33}^{Q} * 1e12$	0.0001	0.0158	0.0000
b_{21}	30175.4829	412859.4780	0.0731
b_{31}	29118.9252	340783.4275	0.0854
ρ_o	0.0892	0.0150	**5.9397**
ρ_1	0.0015	0.0002	**6.2776**
ρ_2	0.0000	0.0001	0.1478
$\rho_3 * 1e9$	0.2243	0.0000	0.0003
k_{o1}^{P}	2.8781	111759.6540	0.0000
k_{o2}^{P}	-0.0011	311.1590	0.0000
k_{o3}^{P}	0.0002	119.5947	0.0000
k_{11}^{P}	0.5517	8.3876	0.0658
k_{21}^{P}	741.5420	5173.3632	0.1433
k_{22}^{P}	1.1635	0.2849	**4.0839**
$k_{23}^{P} * 1e3$	0.0139	0.5860	0.0000
k_{31}^{P}	-0.0068	101.0657	-0.0001
k_{32}^{P}	-0.0003	0.1738	-0.0017
k_{33}^{P}	0.3992	0.3730	1.0702

Table C.1: Point parameter estimates for the $A_1(3)$ benchmark model. Estimates are calculated via the simulated maximum likelihood technique described in section 3.3.1. Sample period is from January 1997 to January 2007, including 521 weekly samples. Standard errors are computed via the product of outer gradients [3].

Parameter	Estimate	Standard Error	t-Statistic
k_{o1}^{Q}	628.3403	3401.5785	0.1847
k_{11}^{Q}	1.2650	0.0929	**13.6216**
k_{21}^{Q}	0.2270	1.8885	0.1202
k_{22}^{Q}	1.6641	0.1129	**14.7442**
k_{23}^{Q}	-0.2510	7.3466	-0.0342
k_{31}^{Q}	-0.0007	0.0220	-0.0309
k_{32}^{Q}	0.0382	1.1084	0.0345
k_{33}^{Q}	0.0001	0.0106	0.0113
b_{21}	0.0090	0.1898	0.0472
b_{31}	0.0314	2.0326	0.0155
ρ_o	-0.1203	0.3153	-0.3816
ρ_1	0.0001	0.0002	0.3639
$\rho_2 * 1e6$	0.0042	0.0000	0.0025
ρ_3	0.0002	0.0068	0.0313
λ_o	2.7178	803.1086	0.0034
λ_θ	-12404.1132	104.4304	**-118.7788**
λ_1	4.6671	12.6287	0.3696
λ_2	201.5465	2044.9756	0.0986
λ_3	-0.0004	0.0690	-0.0056
k_{o1}^{P}	2134.1209	11554.6181	0.1847
k_{o2}^{P}	-0.0133	215.3692	-0.0001
k_{o3}^{P}	399.2485	13461.5025	0.0297
k_{11}^{P}	4.3172	0.9063	**4.7633**
k_{21}^{P}	0.0187	0.2171	0.0859
k_{22}^{P}	0.8700	0.4686	**1.9766**
k_{23}^{P}	-0.0027	0.4427	-0.0060
k_{31}^{P}	0.1825	5.5741	0.0327
k_{32}^{P}	0.1272	3.8043	0.0334
k_{33}^{P}	0.7886	0.7130	1.1061

Table C.2: Point parameter estimates for the target model with information content of the full term structure of target rates incorporated. Estimates are calculated via the simulated maximum likelihood technique described in section 3.3.1. Sample period is from January 1997 to January 2007, including 521 weekly samples. Standard errors are computed via the product of outer gradients [3].

Maturity (yr)	RMSE via map (bp)	RMSE via Particle Filter (bp)
0.5	0	3.7119
1.0	5.0916	3.3423
2.0	0	2.5893
3.0	3.3994	1.4511
4.0	4.7618	1.4019
5.0	4.8754	1.8311
6.0	4.2482	1.6510
7.0	3.3272	1.1959
8.0	2.3245	0.8059
9.0	1.2842	1.1955
10.0	0	2.0592

Table C.3: In sample pricing errors for the target model with information content of the full term structure of target rates incorporated. RMSE via map is found by observing the six month, two year, and ten year yields with no error and inverting the measurement equation. RMSE via Particle Filter is obtained by applying the Bayesian Particle Filter of section 3.3.2.

Bibliography

[1] Cbot fed funds futures. *Chicago Board of Trade*, 2003.

[2] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(No. 2), Feb. 2002.

[3] E. Berndt, B. Hall, R. Hall, and J. Hausman. Estimation and inference in nonlinear structure models. *Annals of Economic and Social Measurement*, 3, 1974.

[4] T. Bjork. *Arbitrage Theory in Continuous Time, 2nd Edition.*

[5] E. F. Bliss and R. R. Bliss. The information in long-maturity forward rates. *American Economic Review*, Vol. 77, 1987.

[6] Federal Reserve Board. Federal board of reserve of minneapolis. *minneapolisfed.org/info/policy/dates-hist.cfm*, 2006.

[7] J. Y. Campbell and R. J. Shiller. Yield spreads and interest rate movements: A bird's eye view. *The Review of Economic Studies, Special Issue: The Econometrics of Financial*, Vol. 58, No. 3, 1991.

[8] O. Cappe, E. Moulines, and T. Ryden. *Inference in Hidden Markov Models.* Springer, New York, 2005.

[9] J. Carlson, J. McIntire, and J. Thomson. Fed funds futures as an indicator of future monetary policy. *Bederal Reserve Bank of Cleveland, Economic Commentary*, November 2003.

[10] G. Chacko and S. Das. Pricing interest rate derivatives: A general approach. *The Review of Financial Studies*, Vol. 15(No. 1), Spring 2002.

[11] Z. Chen. Bayesian filtering: From kalman filters to particle filters, and beyond. *Unpublished Manuscript McMaster Univeristy*, 2003.

[12] P. Cheridito, D. Filipovic, and R.L. Kimmel. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83(1), 2007.

[13] S. V. Chernenko, K. B. Schwarz, and J. H. Wright. The information content of forward and futures prices: Market expectations and the price of risk. *Board of Governors of the Federal Reserve System: International Finance Disscussion Papers*, (8808), 2004.

[14] J. H. Cochrane and M. Piazzesi. Bond risk premia. *American Economic Review*, Vol. 95, 2005.

[15] R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman & Hall/CRC, 2004.

[16] T. Cook and T. Hahn. The effect of changes in federal funds rate target on market interest rates in the 1970s. *Journal of Monetary Economics*, (Vol. 24), 1989.

[17] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interset rates. *Econometrica*, Vol. 53(No. 2), March 1985.

[18] Q. Dai and K. Singleton. Term structure dynamics in theory and reality. *The Review of Financial Studies*, Vol. 16(No. 3), 2003.

[19] Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, Vol. 55(No. 5), 2000.

[20] Q. Dai and K. J. Singleton. Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics*, Vol. 63, 2002.

[21] Datastream. Datastream: Data guide 9.0. *DataStream User Guide Series*, 2002.

[22] A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice.* Springer-Verlag, New York, 2001.

[23] Jefferson Duarte. Evaluating an alternative risk preference in affine term structure models. *The Review of Financial Studies*, 17(2), 2004.

[24] G. R. Duffee. Term premia and interest rate forecasts in affine models. *The Journal of Finance*, Vol. 57, No. 1.

[25] D. Duffie. *Dynamic Asset Pricing Theory.* Princeton University Press, 2001.

[26] D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and application in finance.

[27] D. Duffie and R. Kan. Multi-factor term structure models. *Philosophical Transactions: Physical Sciences and Engineering*, Vol. 347(No. 1684), 1994.

[28] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6, 1996.

[29] D. Duffie, J. Pan, and K. J. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, Vol. 68(No. 6), Nov.

[30] G. Durham and A. R. Gallant. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business and Economic Statistics*, 20(No. 3), July 2002.

[31] C. Edwards and G. Sinzdak. Open market operations in the 1990s. *Bederal Reserve Bank of New York, Economic Commentary*, http://www.federalreserve.gov/pubs/bulletin/1997/199711lead.pdf, 1997.

[32] O. Elerian, C. Siddhartha, and N. Shephard. Likelihood inference for discretely observed nonlinear diffusions. *Econometria*, Vol. 69(No. 4), July 2001.

[33] M. J. Fleming and M. Piazzesi. Monetary policy tick-by-tick. *Federal Reserve Bank of New York*, August 2005.

[34] M. J. Fleming and E. M Remolona. What moves the bond market? *Federal Reserve Bank of New York*, Economic Policy Review, December, 1997.

[35] J. Geweke. Bayesian inference in econometric models using monte carlo integration. *Econometrica*, 57(No. 6), Nov. 1989.

[36] M. Goodfriend and W. Whelpley. *Federal Funds.* Federal Reserve Bank of Richmond, 1993.

[37] N. Gordon, D. Salmond, and A.F.M Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceddings*, 1993.

[38] M. Harrison and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20, 1979.

[39] M. Harrison and S. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, 11, 1981.

[40] M. Heidari and L. Wu. Market anticipation of fed policy change sand the term structure of interest rates. Working Paper, March 2006.

[41] J. James and N. Webber. *Interest Rate Modelling: Financial Engineering.* Wiley, 2000.

[42] G. Jiang and S. Yan. Linear-quadratic term structure models - toward the understanding of jump in interst rates. *Journal of Banking & Finance*, (10.1016), 2008.

[43] M. Johannes. The statisitcal and economic role of jumps in continuous-time interest rate models. *The Journal of Finance*, Vol. LIX, No. 1, 2004.

[44] M. Johannes and N. Polson. *MCMC Methods for Continuous-Time Financial Econometrics.* forthcoming.

[45] K. N. Kuttner. Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics*, 47, 2001.

[46] R. Litterman and J. Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, (54-61), 1991.

[47] J. Lund. Econometric analysis of continuous-time models of the term structure of interest rates. *The Aarhus School of Business*, Working Paper, 1994.

[48] B. Okensdal. *Stochastic Differential Equations, Ed. 6.* Springer–Verlag, Berlin–Heidelberg–New York, 2003.

[49] A. R. Pedersen. A new apporach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scand. J. Statistics*, 1995.

[50] K. B. Petersen and V. Pozdnyakov. Predicting the fed. *Univesity of Connecticut, Department of Economics*, Working Paper, 2008.

[51] M. Piazzesi. Bond yields and the federal reserve. *Journal of Political Economy*, Vol. 113(No. 2), 2005.

[52] M. Piazzesi. *Affine Term Structure Models.* unpublished, 2006.

[53] M. Piazzesi and E. Swanson. Futures prices as risk-adjusted forecasts of monetary policy. *National Bureau of Economic Research*, (Working Paper 10547), 2004.

[54] M.K. Pitt and N. Shepard. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, Vol. 94, 1999.

[55] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods, 2nd Edition.* Springer, New York, 2004.

[56] W. J. Runggaldier. Jump-diffusions models. *Universita di Padova*, Working Paper, 2002.

[57] K. J. Singleton. *Empirical Dynamic Asset Pricing.* Princeton University Press, 2006.

[58] K. J. Singleton. *Empirical dynamic asset pricing :model specification and econometric assessment.* Princeton University Press, 2006.

[59] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, Vol. 5, 1977.