Bond Yields and the Federal Reserve

Monika Piazzesi

University of Chicago and National Bureau of Economic Research

Bond yields respond to policy decisions by the Federal Reserve and vice versa. To learn about these responses, I model a high-frequency policy rule based on yield curve information and an arbitrage-free bond market. In continuous time, the Fed's target is a pure jump process. Jump intensities depend on the state of the economy and the meeting calendar of the Federal Open Market Committee. The model has closed-form solutions for yields as functions of a few state variables. Introducing monetary policy helps to match the whole yield curve, because the target is an observable state variable that pins down its short end and introduces important seasonalities around FOMC meetings. The volatility of yields is "snake shaped," which the model explains with policy inertia. The policy rule crucially depends on the two-year yield and describes Fed policy better than Taylor rules.

This paper is based on chap. 4 of my Stanford PhD dissertation. I am still looking for words that express my gratitude to Darrell Duffie. I would like to thank Andrew Ang, Michael Brandt, John Cochrane, Heber Farnsworth, Silverio Foresi, Lars Hansen, Ken Judd, Tom Sargent, Ken Singleton, John Shoven, John Taylor, and Harald Uhlig for helpful suggestions and Martin Schneider for extensive discussions. I am also grateful for comments from two referees and many seminar participants at Berkeley, the Bank for International Settlements, Carnegie Mellon, Chicago, Columbia, Cornell, the European Central Bank, Harvard, London Business School, London School of Economics, Massachusetts Institute of Technology, the NBER spring 2000 Asset Pricing meeting, the NBER 2000 Summer Institute, Northwestern, the New York Federal Reserve, New York University, Princeton, Rochester, Stanford, Tel Aviv, Tilburg, Toulouse, University of British Columbia, University College London, University of California at Los Angeles, University of Southern California, the 2000 meeting of the Western Finance Association, the 2000 Workshop of Mathematical Finance at Stanford, and Yale. The financial support of doctoral fellowships from the Bradley and Alfred P. Sloan Foundations is gratefully acknowledged.

[*Journal of Political Economy,* 2005, vol. 113, no. 2]
© 2005 by The University of Chicago. All rights reserved. 0022-3808/2005/11302-0005$10.00

I. Introduction

Meeting days of the Federal Open Market Committee (FOMC) are marked as special events on the calendars of many market participants. FOMC announcements often cause strong reactions in bond and stock markets. Indeed, a large literature on announcement effects has documented increased volatility of interest rates at all maturities, not only on FOMC meeting days but also around releases of key macroeconomic aggregates. Not only do markets watch the Federal Reserve, but the reverse is also true. At its meetings, the FOMC extracts information about the state of the economy from the current yield curve. This yield-based information may underlie the FOMC's policy decisions.

These observations suggest that models of the yield curve should take into account monetary policy actions by the Federal Reserve. The extensive term structure literature in finance, however, builds models around a few unobservable state variables, or latent factors, which are backed out from yield data. This statistical description of yields offers only limited insights into the nature of the shocks that drive yields. Moreover, the fit of these models for yields with maturities far away from those included in the estimation is typically bad. This is especially true for short maturities, because most studies avoid dealing with the extreme volatility and the large outliers at certain calendar days of short-rate data.

The above observations also suggest that vector autoregressions (VARs) in macroeconomics that try to disentangle exogenous policy shocks from systematic responses of the Federal Reserve to changes in macroeconomic conditions should take into account yield data. Financial market information, however, is usually not included in VARs, presumably because the usual recursive identification scheme does not work with monthly or quarterly data. Does the Fed not react to current yield data or do yields not react to current policy actions? Each FOMC meeting starts with a review of the "financial outlook," which excludes the first option.[1] And financial markets immediately react to FOMC announcements, which excludes the second.[2]

This paper attempts to kill these two birds with one stone. With high-frequency data, I can use information about the exact timing of FOMC meetings to improve bond pricing and to identify monetary policy shocks. I therefore construct a continuous-time model of the joint dis-

[1] Meyer (1998) takes a very interesting look inside these meetings.

[2] For an excellent survey, see Christiano, Eichenbaum, and Evans (1999). Evans and Marshall (1998) include long yields in a VAR and assume that the Fed does not take into account any information contained in these yields, current or lagged. Eichenbaum and Evans (1995) assume that the Fed conditions on exchange rates from last quarter and ignores more recent exchange rate data. Bagliano and Favero (1998) assume that yields do not react to current policy shocks.

tribution of bond yields and the interest rate target set by the FOMC. The model imposes no arbitrage and respects the timing of FOMC meetings. Decisions about target moves are made at points in time, resulting in a series of target values that looks like a pure jump process. The arrival intensity of target jumps depends on the FOMC meeting calendar and the state of the economy. The model has closed-form solutions for bond prices, which are functions of a small number of state variables.

Closed-form solutions open the door to estimation methods that exploit data on the entire cross section of yields as opposed to a single short rate. Longer yields have the statistical advantage of providing important additional observations, especially in the context of rare policy events. Long yields also have an economic advantage, because they turn out to be inputs in the Fed's policy rule—its systematic response to the state of the economy.

To identify the rule, I rely on the fact that the policy decision is based on information available right before the FOMC starts its meeting. This short informational lag provides a recursive identification scheme. The scheme turns the target forecast from right before the FOMC meeting into a *high-frequency policy rule* and the associated forecast errors into policy shocks. To see what we can learn from the arbitrage-free yield curve model together with this new identifying assumption, I estimate the model with data on short London Interbank Offered Rate (LIBOR) and long swap yields. The model is estimated by the method of simulated maximum likelihood (Pedersen 1995; Santa-Clara 1995), which I extend to jumps.

There are four main estimation results. First, the model considerably improves the performance of existing yield curve models with three latent factors (such as Dai and Singleton [2000]), especially at the short end of the yield curve. Intuitively, the target set by the Fed is an *observable factor* in the model and provides a clean measure of the short end of the yield curve. The use of target data avoids having to deal with calendar day effects in very short rates, which typically require lots of parameters. For example, Hamilton (1996) and Balduzzi, Bertola, and Foresi (1997) use dummies in the mean and variance of the federal funds rate for each day in the reserve maintenance period. These seasonalities, however, do not affect longer yields. For the purpose of modeling the whole yield curve, they can therefore be thought of as seasonal measurement errors. Of course, target data are also affected by seasonalities, those introduced by the FOMC meeting calendar. But the empirical results in this paper suggest that FOMC meetings affect the whole curve and are therefore important for yield curve modeling.

Second, the estimated response of yields to policy shocks is strong and slowly declines only with the maturity of the yield. This response

is roughly consistent with regression results by Cochrane (1989), Evans and Marshall (1998), and Kuttner (2001).

Third, the estimated policy rule describes the Fed as reacting to information contained in the yield curve. I find that the most important information is contained in *yields with maturities around two years*, which suggests that the Fed reacts to some medium-run forecast of the economy. The estimated policy rule displays interest rate smoothing: the target level is autocorrelated. The rule also displays policy inertia: the Fed only partially adjusts the target to its desired rate. Inertia leads to positive autocorrelation in target changes, because one change is typically followed by additional changes in the same direction over a number of FOMC meetings.

As a description of target dynamics, the estimated policy rule performs better than several benchmarks, including estimated versions of the Taylor rule (Taylor 1993). The reason is that yield data summarize market expectations of future target moves. These market expectations are based on a host of variables that are omitted from other rules. Also, yield data are available at higher frequencies and are less affected by measurement errors than macroeconomic variables.

Fourth, I document a *snake shape of the volatility curve*, the standard deviation of yield changes as a function of maturity. Volatility is high for very short maturities (the head of the snake), rapidly decreases until maturities of around three months (the neck of the snake), then increases until maturities of up to two years (the back of the snake), and finally decreases again (its tail). The model explains this snake shape, especially the back of the snake (already documented in Amin and Morton [1994]), with inertia in monetary policy. I also document a calendar effect in the volatility curve around FOMC meetings. The volatility curve shifts up around these meetings, especially at short maturities. The model matches this seasonality with monetary policy shocks, which happen mostly at these meetings.

Related literature.—Papers on yield curve models back out low-dimensional state vectors from yield data. Piazzesi (2004) provides a survey of these models. To capture FOMC decisions, I use a model in the affine class (Duffie and Kan 1996). Most empirical applications treat the factors as latent (among others, Dai and Singleton [2000]), whereas the target is an observable factor in this paper. Few papers in the term structure literature capture aspects of monetary policy. Babbs and Webber (1993) and Farnsworth and Bass (2003) write down theoretical models that do not have tractable solutions for yields. Therefore, they do not take these models to the data.

Most empirical papers on monetary policy focus on the short-rate process alone (Das 2002; Hamilton and Jorda 2002; Johannes 2004). A couple of papers estimate the short-rate process using data on short

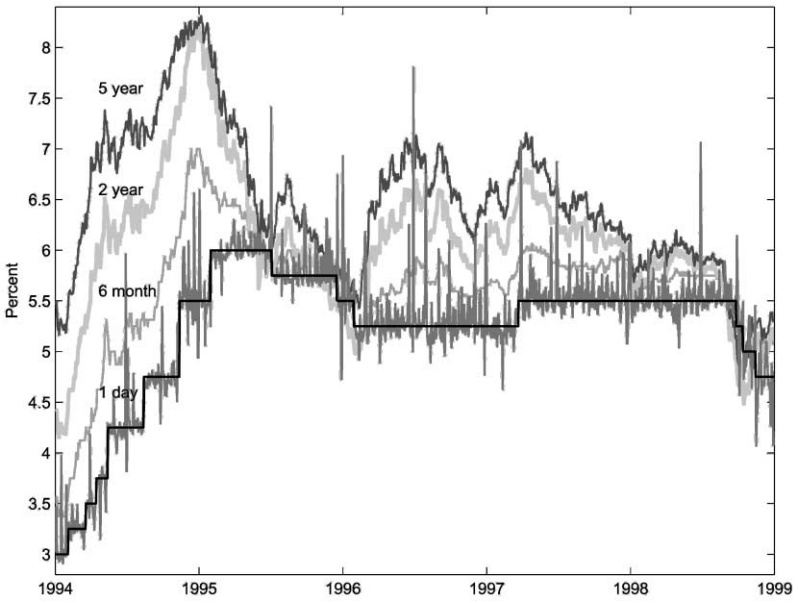


FIG. 1.—Daily data on target (step function), federal funds rate (one-day), LIBOR (six-month), and swap yields (two- and five-year), 1994–98.

rates and then compute long yields using the expectations hypothesis (Rudebusch 1995; Balduzzi et al. 1997). These models cannot match the long end of the yield curve, because the estimation involves only short-end data. Also, there is strong evidence against the expectations hypothesis (Fama and Bliss 1987; Campbell and Shiller 1991). Finally, these papers are not interested in the Fed's policy rule. Kuttner (2001) and others use federal funds futures data and again the expectations hypothesis to define an expected target.

II. FOMC Decisions after 1994

The Federal Reserve targets the overnight rate in the federal funds market. The FOMC fixes a value for the target and communicates it to the Trading Desk of the Federal Reserve Bank of New York, which then implements it through open-market operations (Meulendyke 1998). Figure 1 plots the federal funds target together with LIBOR and swap rates from 1994 to 1998. (Section IV.*A* provides a description of the target data used in this paper.) Looking at the figure, we can see two important stylized facts about Fed targeting. First, the level of the target is persistent. This fact is usually referred to as interest rate smoothing by the

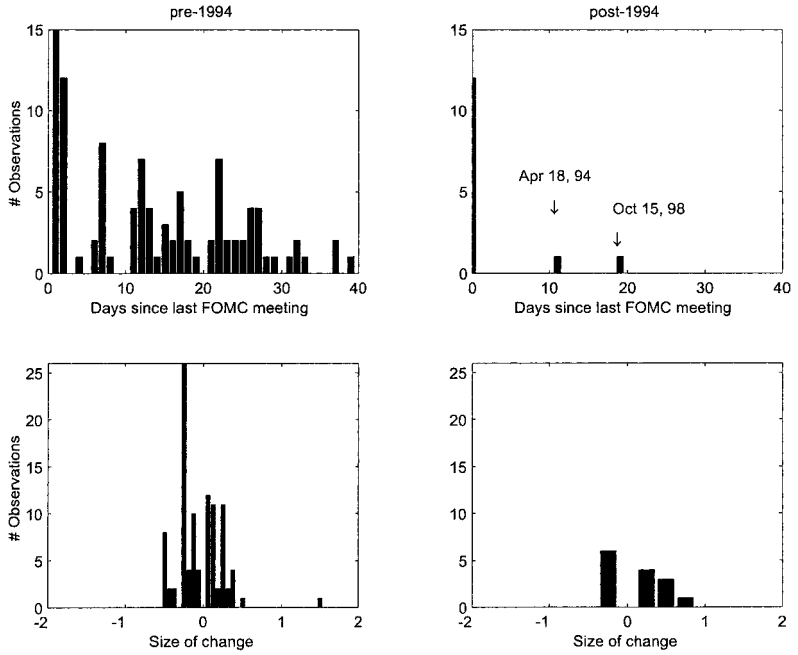


FIG. 2.—The graphs in the first row show the histogram of days since the last FOMC meeting for any given target change between 1984–93 and 1994–98. In the first subperiod, there have been a total of 100 target moves, and there were 14 in the second subperiod. The graphs in the second row show the histogram of the size of target changes for the two subsamples.

Fed. Second, target changes are often followed by additional changes in the same direction. This second stylized fact is called policy inertia.

In 1994, the Fed drastically changed its operating procedures. This change underlies the choice of sample period in this paper, which focuses on the policy framework in place today. Starting with the first FOMC meeting of 1994, the Fed has been announcing the new target at the end of each meeting. The Fed also changed the size and timing of target moves. These latter changes in operating procedures can be seen from figure 2.

The upper row of graphs consists of two histograms, pre-1994 and post-1994, of the number of days between a target change and the preceding FOMC meeting. If, in any given subperiod, the Fed had moved its target only at FOMC meetings, there would be a single spike at 0 in the corresponding histogram. One sees a definite change in 1994 of retargeting mainly at FOMC meeting days, with two exceptions (April 18, 1994, and October 15, 1998) during the data sample used in this analysis and three more exceptions (January 3, April 18, and September

17, 2001) after the end of the sample. The lower row of graphs in figure 2 shows the histogram of target changes for the two subperiods. While pre-1994 target rate changes came in multiples of 6.25 basis points (0.0625 percentage points), after 1994 the Fed used multiples of quarter percentage points.

Under the new operating procedures, "Fed watching" has become a different game. The FOMC meeting calendar has become very important, and investors make forecasts for upcoming meetings. These forecasts are based on a wealth of information including macroeconomic variables (such as consumer prices, gross domestic product, etc.) and even statements by Fed officials themselves (such as U.S. Senate testimonies by the Fed chairman). Also, during a brief time period in 1999, the Fed experimented with announcing its bias regarding future decisions along with its current target decision. Any of this information about future FOMC decisions will be reflected in bond yields, which are used to back out the latent variables in the model. The conditional probability of a target move at upcoming FOMC meetings depends on these latent variables and therefore reflects this information.

The exact timing of intermeeting moves is difficult, if not impossible, to predict. For some of these moves, we know the event that triggered them, such as the Russian financial crisis or the terrorist attacks on September 11. For other moves, it is even difficult to pinpoint the event that triggered them. For example, some say that high car sales in March 1994 suddenly shifted the Fed's assessment of market conditions. Others say that the April 18 move was just a manifestation of authority by Alan Greenspan, because no vote was held. These examples illustrate that it makes sense to assign a small and constant probability to a target move on any given business day.

III. Yield Curve Model with FOMC Decisions

A. *Model*

The state vector $\boldsymbol{X}$ is $\boldsymbol{X}_t = [\theta_t\ s_t\ v_t\ z_t]^\top$, where θ is the federal funds target, $s = r - \theta$ is the spread between the short rate and the target, v is the volatility of s, and z captures other macroeconomic information the Fed uses in setting the target. All variables except θ are unobservable but can be inferred from yields through the bond-pricing model. The dynamics of the state variables are

$$d\theta_t = 0.0025(dN_t^U - dN_t^D), \tag{1}$$

$$ds_t = -\kappa_s s_t dt + \sqrt{v_t} dw_t^s, \tag{2}$$

$$dv_t = \kappa_v(\bar{v} - v_t)dt + \sigma_v\sqrt{v_t}dw_t^v, \tag{3}$$

$$dz_t = -\kappa_z z_t dt + dw_t^z, \tag{4}$$

where N^U and N^D are counting processes with stochastic intensities λ^U and λ^D, respectively, and w^s, w^v, and w^z are independent Brownian motions.

Now I describe the state variables in more detail. Following the usual convention, one year is an interval of length one, and yields are annualized percentages (0.05 is 5 percent). Figure 1 shows that sample paths of the target are step functions. The steps are multiples of 25 basis points (bp), or 0.0025. In continuous time, the target is a pure jump process given by (1). Target jumps up and down are counted by N^U and N^D, respectively. Heuristically, the probability of a jump in N^U during the interval $[t, t + dt]$ conditional on information up to time t is given by $\lambda_t^U dt$, and the conditional probability of a jump in N^D is $\lambda_t^D dt$. The conditional probability of, say, a target increase by 25 bp during $[t, t + dt]$ is then $\lambda_t^U dt \times (1 - \lambda_t^D dt)$. The econometrician has discrete observations only on the difference between N^U and N^D. This means that the econometrician gets to observe target moves of 0 bp, ± 25 bp, ± 50 bp, and so forth.

Figure 1 shows large spikes in the federal funds rate around certain calendar dates, such as the end of the year or so-called settlement Wednesdays. I treat these spikes as seasonal measurement errors. In other words, these seasonalities do not affect the true short rate r or the dynamics (2) of the spread s. Figure 1 also suggests that the short rate reverts back to the target. The spread dynamics therefore pull s back to zero at speed κ_s. To capture fat tails in the yield distribution between FOMC meetings, I use stochastic volatility (3). The parameter $\bar{v}$ is the mean volatility, κ_v is the speed of mean reversion, and σ_v controls the size of shocks to v.

By far the most interesting state variable is z. The process z enters the model only through its influence on the jump intensities λ^U and λ^D, which will be specified below. Its value z_t at time t proxies for macro information the Fed cares about when setting the target—information that is not already contained in the other state variables. The model implies a solution for yields at time t as a function of $\mathbf{X}_t$, so that this information can be backed out from yield data. The process (4) has mean zero and is normally distributed.

1. Probability of Target Moves

The Fed sets the target in response to the value of $\mathbf{X}$. The conditional probability of a target move varies according to the FOMC meeting

calendar. Outside of FOMC meetings, there is a small and constant probability of a move. FOMC meetings are time intervals; the ith meeting is $[t_{i-1}, t_i]$. During any such interval, the intensities take the form

$$\lambda_t^U = \overline{\lambda} + \boldsymbol{\lambda}_X^\top(\boldsymbol{X}_t - \overline{\boldsymbol{X}}),$$
$$\lambda_t^D = \overline{\lambda} - \boldsymbol{\lambda}_X^\top(\boldsymbol{X}_t - \overline{\boldsymbol{X}}), \qquad t \in [t_{i-1}, t_i]. \tag{5}$$

These intensities depend on the distance of $\boldsymbol{X}_t$ from its mean $\overline{\boldsymbol{X}}$. The intensities are, on average, equal to $\overline{\lambda}$, and the parameters in $\boldsymbol{\lambda}_X \in \mathbb{R}^N$ control their time variation. The plus and minus signs in front of $\boldsymbol{\lambda}_X$ make the intensities move in opposite ways over time. I shall deal with negative values for intensities and target in Section IV.*D*.

2. Identification of the High-Frequency Policy Rule

To identify a structural equation that describes the Fed's behavior, I assume that the Fed reacts to information "right before" the FOMC meeting. This is a natural assumption: FOMC members meet and discuss data available up to that time, including bond market data, but not yield changes during the meeting. The assumption amounts to a recursive identification scheme. The scheme turns the expected value of the new target conditional on the value of $\boldsymbol{X}$ at the beginning of the meeting into a high-frequency policy rule, whereas unexpected target changes are identified as policy shocks.

To write down the rule, I define monetary policy shocks $M = M^U - M^D$, where M^j is the compensated process $\{M_t^j = N_t^j - \int_0^t \lambda_u^j du;\ t \geq 0\}$ for $j = U, D$. Heuristically, the conditional expected value of dN_t^j is $1 \times \lambda_t^j dt$, which implies that dM_t^j and dM_t are mean zero shock series with a nonnormal distribution. Now I can write

$$d\theta_t = E_t[d\theta_t] + 0.0025\, dM_t, \tag{6}$$

where the expected target change during an FOMC meeting[3] is

$$E_t[d\theta_t] = -2\boldsymbol{\lambda}_X^\top(\overline{\boldsymbol{X}} - \boldsymbol{X}_t)dt = \kappa_\theta[(a + \boldsymbol{b}^\top \boldsymbol{X}_t) - \theta_t]dt. \tag{7}$$

The second equality introduces the scalars κ_θ, a, and the parameters in $\boldsymbol{b} \in \mathbb{R}^N$. The last term can be interpreted as a partial adjustment of the current target θ_t to a desired rate $a + \boldsymbol{b}^\top \boldsymbol{X}_t$. The speed of this adjustment is κ_θ. To get the policy rule, we need to sum up expected target changes (7) during an FOMC meeting and apply the law of iterated expectations.

[3] I fix the arrival rates of target moves outside of FOMC meetings to their empirical frequency. There has been one up and one down move outside of FOMC meetings during the five years from 1994 to 1998, so I set $\lambda_t^U = \lambda_t^D = 0.2$ outside of meetings. This implies that $E_t[d\theta_t] = 0$ outside of FOMC meetings.

3. Pricing Kernel

The pricing kernel is the product of marginal utility divided by the price of consumption. I do not specify preferences together with processes for consumption and prices. Instead, I specify the pricing kernel directly as a function of state variables:

$$M_t = \exp\left(-\int_0^t r_s ds\right)\xi_t, \tag{8}$$

where

$$\frac{d\xi_t}{\xi_t} = -\boldsymbol{\sigma}_\xi(\boldsymbol{X}_t)^\top d\boldsymbol{w}_t, \tag{9}$$

and $\boldsymbol{w} = [0\ w^s\ w^v\ w^z]^\top$. The vector $\boldsymbol{\sigma}_\xi(\boldsymbol{X}_t)$ contains the market prices of risk for the various Brownian motions. I assume that it has the form

$$\boldsymbol{\sigma}_\xi(\boldsymbol{X}_t) = [0\ q_s\sqrt{v_t}\ q_v\sigma_v\sqrt{v_t}\ q_z]^\top, \tag{10}$$

where q_s, q_v, and q_z are constants (as in Longstaff and Schwartz [1992]). I do not allow the pricing kernel to jump.[4]

An alternative interpretation, which I shall refer to below, is to use the process ξ as density to define a probability measure Q, under which risk-neutral pricing applies. The expectation under the risk-neutral measure satisfies $E_t[Y\xi_s/\xi_t] = E_t^Q[Y]$ for any random variable Y known at time $s \geq t$ for which this expectation exists. Under Q, a standard Brownian motion $\boldsymbol{w}^Q$ solves $d\boldsymbol{w}_t^Q = d\boldsymbol{w}_t + \boldsymbol{\sigma}_\xi(\boldsymbol{X}_t)dt$. To see the dynamics of the state variables under the risk-neutral measure, we can simply insert $d\boldsymbol{w}_t = d\boldsymbol{w}_t^Q - \boldsymbol{\sigma}_\xi(\boldsymbol{X}_t)dt$ into (1)–(4).

B. *Solving for Yields*

Asset prices are expected future payoffs weighted with the pricing kernel. Equivalently, asset prices are expected discounted payoffs under the risk-neutral probability measure Q. From equation (8), the price $P(t, T)$ at time t of a zero-coupon bond that pays \$1 at time T is

$$\begin{aligned} P(t, T) &= E_t\left[\frac{M_T}{M_t}\right] = E_t\left[\frac{\xi_T}{\xi_t}\exp\left(-\int_t^T r_u du\right)\right] \\ &= E_t^Q\left[\exp\left(-\int_t^T r_u du\right)\right]. \end{aligned} \tag{11}$$

[4] Since the sample is short, the estimation of jump parameters for ξ is difficult. For example, $\overline{\lambda}$ is estimated imprecisely even in the absence of jump risk prices.

The short rate r is the sum of θ and s, which solve stochastic differential equations (1) and (2), respectively.

The solution to (11) satisfies a partial differential integral equation (PDIE) stated in Appendix A. The solution to this PDIE is an exponential affine function in the state variables:

$$P(t,\ T) = \exp\left[\bar{c}(t,\ T) + c_\theta(t,\ T)\theta_t + c_s(t,\ T)s_t + c_v(t,\ T)v_t + c_z(t,\ T)z_t\right] \tag{12}$$

for coefficients $\bar{c}(t,\ T)$ and

$$\boldsymbol{c}_{\boldsymbol{X}}(t,\ T) = [c_\theta(t,\ T)\ \ c_s(t,\ T)\ \ c_v(t,\ T)\ \ c_z(t,\ T)]^\top$$

that solve ordinary differential equations (ODEs). The ODEs are stated in Appendix A. Zero-coupon yields are linear:

$$Y_0(t,\ T) = -\frac{\ln P(t,\ T)}{T-t} = \bar{c}^y(t,\ T) + \boldsymbol{c}_{\boldsymbol{X}}^y(t,\ T)^\top \boldsymbol{X}_t, \tag{13}$$

with $\bar{c}^y(t,\ T) = -\bar{c}(t,\ T)\ /\ (T-t)$ and $\boldsymbol{c}_{\boldsymbol{X}}^y(t,\ T) = -\boldsymbol{c}_{\boldsymbol{X}}(t,\ T)/(T-t)$.

Most models have yield coefficients that depend only on time to maturity $T-t$. By contrast, the yield coefficients $\bar{c}^y(t,\ T)$ and $\boldsymbol{c}_{\boldsymbol{X}}^y(t,\ T)$ in this model depend on the particular ordering of FOMC meetings between t and T and therefore on t and T separately. I therefore cannot follow the usual procedure of computing the yield coefficients as a function of $T-t$ by starting at zero time to maturity and solving the ODEs forward. Instead, I need to compute the coefficients for each observation t in the sample and each yield maturity T in the data set separately. This immensely increases the computational burden when evaluating the likelihood function for a candidate parameter value, especially with long yields in the data set. Fortunately, the following algorithm works and saves time. The algorithm matches only the exact number of days until the next FOMC meeting, whereas subsequent meetings are assumed to be equally spaced over the year. This is only an approximation, because the actual calendar time between these FOMC meetings varies. However, the errors due to this approximation are virtually undetectable for the maturities of the yields used in the estimation (six months and above).

The FOMC targets the federal funds rate, which pertains to interbank loans. These loans are not default-free because they are not collateralized. As a result, the federal funds rate and its target are substantially higher than short Treasury-bill rates (which are further depressed by tax and liquidity effects). To estimate the model, I therefore use rates on LIBOR and swap contracts, which are traded mainly between banks. The time t swap rate is the fixed rate at which banks can borrow for τ years in exchange for floating payments that have a discounted present

value of \$1. The contract specifies that both the loan and the floating repayment be paid in biannual installments. The time t swap rate $Y(t, t+\tau)$ is then determined as the rate that equalizes the present discounted value of these installments at time t:

$$1 = P(t,\ t+\tau) + \frac{Y(t,\ t+\tau)}{2}\sum_{j=1}^{2\tau} P(t,\ t+0.5j), \tag{14}$$

where the left-hand side is the \$1 worth of floating repayments and the right-hand side is the value of the biannual fixed loan payments (which explains the division by two). Following Duffie and Singleton (1997), I interpret the symbol r as the rate on short bonds of LIBOR and swap quality. This means that r reflects the credit risk of interbank loans, just like the Fed's target.

IV. Estimation

The parameter vector $\boldsymbol{\gamma}$ contains 14 parameters for the intensities $\overline{\lambda}$, λ_θ, λ_s, λ_v, and λ_z; the mean reversions κ_s, κ_v, and κ_z; the means $\bar{\theta}$ and $\bar{v}$; the volatility σ_v; and the risk premia q_s, q_v, and q_z. For a given parameter vector $\boldsymbol{\gamma}$, the model maps the state vector $\boldsymbol{X}_t$ into observables $\boldsymbol{Y}_t$ based on equation (14). The vector of observables contains the target, the six-month LIBOR, and the two- and five-year swap yields:

$$\boldsymbol{Y}_t = [\theta_t\ Y(t,\ t+0.5)\ \ Y(t,\ t+2)\ \ Y(t,\ t+5)]^\top.$$

Section IV.*A* describes the data on $\boldsymbol{Y}_t$.

Ideally, the parameters would be estimated by maximizing the likelihood function of the observables over $\boldsymbol{\gamma}$. The likelihood function is the product of densities $f(\boldsymbol{Y}_t,\ t|\boldsymbol{Y}_\tau,\ \tau;\ \boldsymbol{\gamma})$ conditional on the last observation Y_τ at some $\tau < t$. The density f can be obtained by a change of variable from the conditional density $f_X(\boldsymbol{X}_t,\ t|\boldsymbol{X}_\tau,\ \tau;\ \boldsymbol{\gamma})$ of $\boldsymbol{X}_t$:

$$f(\boldsymbol{Y}_t,\ t|\boldsymbol{Y}_\tau,\ \tau;\ \boldsymbol{\gamma}) = f_X(\boldsymbol{g}(\boldsymbol{Y}_t,\ \boldsymbol{\gamma}),\ t|\boldsymbol{g}(\boldsymbol{Y}_\tau,\ \boldsymbol{\gamma}),\ \tau;\ \boldsymbol{\gamma})|\nabla_Y \boldsymbol{g}(\boldsymbol{Y}_t,\ \boldsymbol{\gamma})|, \tag{15}$$

where $\boldsymbol{g}(\cdot,\ \boldsymbol{\gamma})$ is the function from the observables to the state vector, in that $\boldsymbol{X}_t = \boldsymbol{g}(\boldsymbol{Y}_t,\ \boldsymbol{\gamma})$. This function inverts the yield formulas (14).

Now, three problems arise. First, the true density f_X is not available in closed form. I therefore extend the simulated maximum likelihood (SML) method of Pedersen (1995) and Santa-Clara (1995) to jump diffusions (Sec. IV.*B*). Second, the function $\boldsymbol{g}(\boldsymbol{Y}_t,\ \boldsymbol{\gamma})$ needs to be inverted numerically for each observation t. To do this, I use a hill-climbing method based on analytical gradients. As a by-product, I get the Jacobian term $|\nabla_Y \boldsymbol{g}(\boldsymbol{Y}_t,\ \boldsymbol{\gamma})|$ analytically (Sec. IV.*C*). Finally, the function $\boldsymbol{g}$ does not impose that intensities and the target need to be positive. To control

the approximation accuracy of $\boldsymbol{g}$, I experiment with constraining the parameter space (Sec. IV.*D*).

A. Data

The sample period is January 1, 1994, to December 31, 1998. The target series is taken from Datastream, except for the timing of the target move in February 1994. Datastream assigns the move to February 3, whereas the move was announced only on February 4 (Bradsher 1994). There are eight FOMC meetings per year. The dates of these meetings come from the Board of Governors of the Federal Reserve. Most meetings are on Tuesdays. Two meetings per year (the first and the fourth) extend over Tuesdays and Wednesdays. For solving yields and setting up their likelihood function, the two-day meetings are dated on Wednesdays, because target decisions are always announced at the end of the meetings.

LIBOR data are taken from the British Bankers' Association, whereas swap rates are taken from Intercapital Brokers Limited. Both series are obtained through Datastream. LIBOR rates are recorded at 11:00 a.m. London time, and swap rates are recorded at the end of the U.K. business day. Target changes are typically announced at 2:15 p.m. Eastern time. These announcements affect swap and LIBOR rates recorded for the next day. There have been a number of exceptions to the 2:15 p.m. rule during 1994 and even after 1994. To make sure that FOMC announcements on Tuesdays or Wednesdays always affect LIBOR and swap rates recorded for the same week, I construct a weekly data set with Thursday (London time) observations of LIBOR and swap yields, together with Wednesday (Eastern time) observations of the target. Whenever the respective day was a holiday, I used the observation of the previous business day.

B. Density Approximation

The conditional density of the state vector solves a partial differential integral equation that has a closed-form solution for only a few special cases, such as Gaussian and square root diffusions. To overcome this problem, I use SML. This estimation method attains approximate efficiency. To fix notation, the state space is $D \subset \mathbb{R}^N$. The conditional density of $\boldsymbol{X}_t$ can be written, using Bayes' rule and the Markov property of $\boldsymbol{X}$, as

$$f_X(\boldsymbol{X}_t,\ t|\boldsymbol{X}_\tau,\ \tau) = \int_D f_X(\boldsymbol{X}_t,\ t|\boldsymbol{x},\ t-h) f_X(\boldsymbol{x},\ t-h|\boldsymbol{X}_\tau,\ \tau) d\boldsymbol{x}, \qquad (16)$$

for any time interval h. (This is called the Chapman-Kolmogorov equation.) SML computes (16) by Monte Carlo integration, replacing the density $f_{\boldsymbol{X}}(\boldsymbol{X}_t, t|\boldsymbol{x}, t-h)$ by the density $\hat{f}_{\boldsymbol{X}}$ of a discretization of $\boldsymbol{X}$.

Appendices B and C explain how to extend SML to jump diffusions. The appendices also explain how to overcome the additional problems associated with estimating the particular model presented in this paper. For example, special care needs to be taken to accommodate stochastic intensities that depend on calendar time. These intensities may become very large to predict multiple target moves during an FOMC meeting. Therefore, the interval h needs to be chosen carefully. Another difficulty is that FOMC meetings may introduce discontinuities in the objective function, when small changes in parameters do not change the number of target moves across simulated samples.

C. SML Likelihood

The SML estimator $\hat{\boldsymbol{\gamma}}$ maximizes the approximate likelihood

$$\prod_{(\tau,t)\in I} \hat{f}(\boldsymbol{Y}_t, t|\boldsymbol{Y}_\tau, \tau; \boldsymbol{\gamma}) = \prod_{(\tau,t)\in I} \hat{f}_{\boldsymbol{X}}(\boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma}), t|\boldsymbol{g}(\boldsymbol{Y}_\tau, \boldsymbol{\gamma}), \tau; \boldsymbol{\gamma})|\nabla_{\boldsymbol{Y}}\boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma})|, \tag{17}$$

where I denotes pairs of successive observation times in the data set.

The mapping $\boldsymbol{g}(\cdot, \boldsymbol{\gamma})$ from observables $\boldsymbol{Y}_t$ to state variables $\boldsymbol{X}_t$ cannot be inverted analytically. The reason is that the swap yield formula (14) is nonlinear. To invert $\boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma})$ numerically for every observation t, I use a hill-climbing procedure. To save time, the procedure uses analytical derivatives:

$$\frac{dY(t, t+\tau)}{d\boldsymbol{X}_t} = -\frac{2\boldsymbol{c}_{\boldsymbol{X}}(t, \tau)P(t, t+\tau) + Y(t, t+\tau)\sum_{j=1}^{2\tau}\boldsymbol{c}_{\boldsymbol{X}}(t, t+0.5j)P(t, t+0.5j)}{\sum_{j=1}^{2\tau}P(t, t+0.5j)}.$$

The 4×4 Jacobian matrix $d\boldsymbol{Y}_t^\top/d\boldsymbol{X}_t$ contains these derivatives for $\tau = 0.5$, two, and five years in its last three columns. Its first column is $d\theta_t/d\boldsymbol{X}_t = [1\ \boldsymbol{0}_{3\times 1}]^\top$, where $\boldsymbol{0}_{3\times 1}$ denotes a 3×1 vector of zeros. To get the Jacobian term for the density, I compute

$$|\nabla_{\boldsymbol{Y}}\boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma})| = \frac{1}{|d\boldsymbol{Y}_t^\top/d\boldsymbol{X}_t|}.$$

D. *Approximation Accuracy*

The mapping $\boldsymbol{g}(\cdot, \boldsymbol{\gamma})$ approximates the true mapping of a model, in which intensities and the target are always positive. The accuracy of this approximation may be unacceptable when we replace $\boldsymbol{\gamma}$ with the *unconstrained estimator* $\hat{\boldsymbol{\gamma}}$. I therefore obtain another set of estimates by constraining the parameter space. Here, the space contains only those parameters at which the observations are explained by a state realization $\boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma})$ for which the intensities are positive. Formally, I define the set

$$A := \{x \in D : \overline{\lambda} + \boldsymbol{\lambda}_X^\top(\boldsymbol{x} - \overline{\boldsymbol{X}}) \geq 0,\ \overline{\lambda} - \boldsymbol{\lambda}_X^\top(\boldsymbol{x} - \overline{\boldsymbol{X}}) \geq 0\}.$$

The *constrained estimator* $\hat{\boldsymbol{\gamma}}_c$ solves

$$\max_{\boldsymbol{\gamma}} \prod_{(\tau,t)\in I} \hat{f}(\boldsymbol{Y}_t,\ t|\boldsymbol{Y}_\tau,\ \tau;\ \boldsymbol{\gamma})$$
$$\text{subject to } \boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma}) \in A \text{ for all } t \in I. \tag{18}$$

Appendix D checks the approximation accuracy of $\boldsymbol{g}(\cdot, \hat{\boldsymbol{\gamma}})$ and $\boldsymbol{g}(\cdot, \hat{\boldsymbol{\gamma}}_c)$. This is done by computing the true function from factors to yields with Monte Carlo methods. The true mapping is then compared to $\boldsymbol{g}$. It turns out that the approximation errors are sufficiently small, for both constrained $\hat{\boldsymbol{\gamma}}_c$ and unconstrained $\hat{\boldsymbol{\gamma}}$ parameter estimates. I shall therefore focus on unconstrained estimates in the rest of this paper.

V. Estimation Results

A. *Parameter Estimates*

Table 1 reports the unconstrained parameter estimates $\hat{\boldsymbol{\gamma}}$ for the model described in equations (1)–(5) and (10). Table 1 also reports the unconstrained estimates for an interesting version of the model that sets volatility of the spread constant: $v_t = \bar{v}$. It is important to note that "constant volatility" here refers only to the spread; the variance of yields still varies over time because of jumps. The constant volatility version is easier to estimate because it has only two latent factors instead of three. To break the resulting singularity, I assume that the two-year swap yield is measured with error. I estimate the autocorrelation coefficient and the variance of this error. Setting $v_t = \bar{v}$ still economizes on parameters because λ_v, σ_v, κ_v, and q_v are not needed.

Table 1 shows that the unconditional probability of a target move up or down is estimated imprecisely; the t-ratio of $\overline{\lambda}$ is below two in both versions of the model. The reason is that the intensities depend on persistent variables, such as the target and stochastic volatility, and the sample is short. To understand the values of the intensity parameters, it is nevertheless useful to look at the point estimate of $\overline{\lambda}$, which is 10.

TABLE 1
SIMULATED MAXIMUM LIKELIHOOD ESTIMATES

	MODEL WITH STOCHASTIC VOLATILITY		MODEL WITH CONSTANT VOLATILITY	
	Estimate	t-Ratio	Estimate	t-Ratio
Mean reversion:				
κ_s	9.75	4.69	1.56	4.10
κ_v	.04	.42	…	…
κ_z	.72	4.34	.29	3.03
Means:				
$\bar{\theta}$	.0522	…	.0522	…
$\bar{v}$	.000415	1.07	…	…
Intensities:				
$\overline{\lambda}$	10	.55	84	.40
λ_θ	−9,408.9	−4.18	−4,876.7	−4.30
λ_s	7,267	1.86	7,582	2.25
λ_v	548,315	1.63	…	…
λ_z	237.6	17.87	119.5	18.63
Risk premia:				
q_s	−47.62	−2.90	70.29	8.67
q_v	−2,537.5	−.81	…	…
q_z	.1126	.18	−.2132	33.36
Volatility:				
σ_v	.0058	.78	…	…
$\sqrt{\bar{v}}$	…	…	.0089	12.50

NOTE.—The model with stochastic volatility refers to eqq. (1)–(5) and (10). The model with constant volatility sets $v_t = \bar{v}$. The estimation of the constant volatility model assumes that the two-year swap yield is measured with error. The autocorrelation of this error is estimated to be 0.955 (with a t-ratio of 15), and its volatility is estimated to be 0.002 (with t-ratio of 56). The parameter $\bar{\theta}$ is fixed to the average target over the sample. Apps. B and C contain practical details about the estimation. The sample is weekly from January 1994 to December 1998.

There are eight FOMC meetings per year, which leads to average intensities of $(\overline{\lambda} \times 8)/365 = 0.22$, roughly one jump every five years. This estimate is low given the five down and seven up moves at FOMC meetings during the five-year sample period, which suggests that $(\overline{\lambda} \times 8)/365$ should be above one. Again, the estimate is too imprecise to hold this against the model. In the constant volatility model, the estimate of $\overline{\lambda}$ is 84, implying 1.8 jumps per year. This point estimate is more reasonable.

The time variation in probabilities is driven by the state variables z and θ. The t-ratios of all other slope parameters in $\lambda_X = [\lambda_\theta \; \lambda_s \; \lambda_z \; \lambda_v]^\top$ are below two. To understand the time variation induced by macro information z, let us first look at the parameter estimates for the z-process itself. The estimated speed of mean reversion in z is $\kappa_z = 0.72$, which amounts to a weekly autoregressive coefficient of $\exp(-\kappa_z/52) = 0.986$ and a half-life of shocks to z of $-\ln(0.5)/\kappa_z = 0.96 \approx 1$ year. In the constant volatility model, the half-life of these macro shocks is estimated to be even longer, around 2.5 years. The estimate of λ_z is positive in both versions of the model, so that a positive shock to z increases

TABLE 2
CORRELATIONS OF STATE VARIABLES, YIELDS, AND TARGET

	$\hat{\gamma}$			LIBOR AND SWAPS			TARGET
	r (1)	z (2)	v (3)	6-Month (4)	2-Year (5)	5-Year (6)	θ (7)
$\hat{\gamma}$:							
r	1			.54	−.03	−.07	−.12
z	−.18	1		.44	.81	.65	.13
v	−.01	.01	1	.37	.55	.76	−.03
$v_t = \bar{v}$:							
r	.67	.24	−.08	.56	.26	.14	.66
z	−.24	.63	.78	.44	.89	.97	.07
Dai-Singleton:							
r	.57	-.90	.05	−.05	−.62	−.50	−.26
θ	−.16	.93	.34	.63	.96	.86	.29
v	.02	−.19	.97	.19	.33	.59	−.21

NOTE.—This table computes the correlation of the first differences of model-implied state variables and data on yields and target. Rows 1–3 use the state variables r, z, and v computed with the estimated parameter values $\hat{\gamma}$ from table 1. Rows 4–5 use the variables r and z from the model with constant volatility $v_t = \bar{v}$. Rows 6–8 use the state variables r, θ, and v from the $A_1(3)_{DS}$ model by Dai and Singleton (2000) computed with their estimated parameters. The correlations are computed over the weekly sample January 1994 to December 1998. The correlations with the target in col. 7 are computed using the subsample of FOMC meetings.

the conditional probability of an up move not only at the next FOMC meeting but also at subsequent meetings. In other words, macro shocks are likely to trigger many target moves in the same direction. Therefore, these shocks induce positive autocorrelation in target changes or *policy inertia*.

To understand what type of information the variable z proxies for, I compute correlations between yield data $\boldsymbol{Y}_t$ and the time series of factors $\boldsymbol{g}(\boldsymbol{Y}_t, \hat{\boldsymbol{\gamma}})$ implied by the yield data at the estimated parameters $\hat{\boldsymbol{\gamma}}$. Rows 1–3 in table 2 report these correlations. Row 2 shows that the macro information in z is closely related to the *two-year swap yield*. Rows 4–5 report correlations for the model with constant volatility $v_t = \bar{v}$. Here, z is closely related to both the two-year and the five-year yields. This is related to the fact that data on longer yields are more persistent and that z is estimated to be more persistent in this version of the model.

The probability of a target move depends on the past target through the parameter λ_θ, which is estimated to be positive. If the target is higher than its mean $\bar{\theta} = 5.22$ percent, there is a high conditional probability of a target cut at the next FOMC meeting (and further meetings down the road). If the target is lower than its mean, there is a high probability of a target increase. Taken together, these effects induce mean reversion. The mean reversion is slow, which captures interest rate smoothing. Deviations of the short rate from the target are estimated to be short-lived; they represent *money market noise*. The speed $\kappa_s = 9.75$ at which shocks to the spread die out implies a weekly autoregressive coefficient

of $\exp(-\kappa_s/52) = 0.83$ and a half-life of shocks to the spread of $-\ln(0.5)/\kappa_s = 0.07$, less than one month.

The short rate $r = \theta + s$ is closely related to other short rates, as we can see from its 54 percent correlation with LIBOR. Rows 6–8 of table 2 report correlations for factors implied by the Dai and Singleton (2000) model. In their model, the short rate is almost uncorrelated with the six-month LIBOR rate; the correlation coefficient is even slightly negative: −5 percent. The short rate in this paper and the Dai-Singleton short rate are thus very different; they are only 57 percent correlated. This difference will be important for the performance of these models when it comes to matching the short end of the yield curve (documented in the next section). At FOMC meetings, the short rate in this model and the Dai-Singleton short rate are both negatively correlated with the target (col. 7 of table 2). Interestingly, the short rate in the constant volatility model does not share this unattractive feature. This short rate is even more closely related to LIBOR and strongly commoves with the target at FOMC meetings. Again, this will show up in performance.

The stochastic volatility factor v is roughly comparable to volatility in the Dai-Singleton model. Table 2 reports that the correlation coefficient between the two variables is 97 percent. Volatility is highly persistent. Its speed of mean reversion, $\kappa_v = 0.04$, in table 1 is close to zero, implying a half-life of shocks of several years. This is also reflected in the high correlation between v and the longest (and thus most autocorrelated) yield in table 2. The parameters related to volatility in table 1 are therefore estimated imprecisely. To keep the number of parameters low, the model assumes that the Brownian motions w^s, w^v, and w^z are orthogonal. As we can see from table 2, this assumption does not seem to miss important correlations in the state variables. The volatility factor v is almost uncorrelated with the short rate and the macroeconomic information contained in z.

B. Bond-Pricing Performance

By construction, the model explains yields used in the estimation without any error. These yields are the six-month LIBOR and the two-year and five-year swap yields. To get a sense of the cross-sectional fit of the model, I look at how the model performs in matching yields not used in the estimation. These are yields with maturities such as one and three months or one, three, and four years. Table 3 reports mean absolute pricing errors for these LIBOR and swap yields. Pricing errors are defined as the difference between actual yields and model-implied yields, which are computed by inserting model-implied factors $\boldsymbol{g}(\boldsymbol{Y}_t, \hat{\boldsymbol{\gamma}})$ into the yield formulas (14). From the first row of table 3, we can see that the four-factor model performs well across all maturities. The model mis-

TABLE 3
PRICING ERRORS (in Basis Points)

	1 Month	3 Months	1 Year	3 Years	4 Years
$\hat{\gamma}$	25.7	11.0	3.2	1.8	1.8
$v_t = \bar{v}$	12.5	7.5	6.8	9.5	5.8
Dai-Singleton	237.3	66.2	9.9	6.4	5.7

NOTE.—This table computes mean absolute pricing errors in basis points over the weekly sample from January 1994 to December 1998 for three different models. The first row uses the model evaluated at the parameter values $\hat{\gamma}$ from table 1. The second row uses a version of the model with constant volatility $v_t = \bar{v}$. The third row shows the mean absolute pricing errors of the $A_1(3)_{DS}$ model by Dai and Singleton (2000) at their parameter estimates.

prices long bonds by only 2 bp. For shorter maturities, the pricing errors are still small, around 26 bp.

The second row of table 3 reports pricing errors with constant volatility: $v_t = \bar{v}$. This version of the model has three factors, but one of the factors is a man-made variable and not a market yield. Therefore, the model is less flexible in matching yields than a model with three latent factors. Despite this, the model performs even better at the short end than the full four-factor model, with pricing errors of only 13 bp. The model makes somewhat larger errors at the long end, but its performance is still pretty good. These results are surprising. Perhaps stochastic volatility is not as important for matching the yield curve, at least not during the last decade. Piazzesi (2004) reports that monthly yield changes have become "more Gaussian," in that they exhibit far less excess kurtosis during the 1990s than during the entire postwar sample.

As a rough benchmark, the third row of table 3 reports pricing errors for the Dai and Singleton (2000) model based on three latent factors. These pricing errors are computed with parameters estimated using different yields (same LIBOR and two-year swap, but a 10-year instead of a five-year swap). They are also based on a sample that only partially overlaps with the sample used in this paper (weekly data from April 1987 to August 1996 instead of January 1994 to December 1998). The benchmark thus serves as only a rough indication of pricing errors rather than as a detailed comparison. Keeping this in mind, we see that the Dai-Singleton model misses the short end of the yield curve by over two percentage points. When this number is compared to the errors made by the three-factor constant volatility model, it seems that it helps to convert one latent factor into the target. In other words, the target appears to fix the short end of the yield curve at a good position. The four-factor model with target performs better than the Dai-Singleton model across the whole curve, but the improvement is most dramatic at the short end.

The source of these performance differences is that the models imply very different short rates. In particular, in this paper the short rate behaves like other short rates, whereas the Dai-Singleton short rate does

not look like any other rate. To start with sample means, the average short-rate series in the four-factor model with target is 5.02 percent, whereas the average Dai-Singleton short rate is −0.46 percent. Table 2 shows that the short rate and LIBOR are 54 percent correlated, whereas the Dai-Singleton short rate is −5 percent correlated with LIBOR and strongly negatively correlated with longer swap yields. As already mentioned, the short rate in the constant volatility model is even more closely related to other short yields, which explains its better performance at the short end of the curve.

C. *High-Frequency Policy Rule*

The high-frequency policy rule is the expected value of the target conditional on information right before the FOMC meeting. More precisely, the rule $E_t[\theta_u]$ is equal to the first component in $E_t[\boldsymbol{X}_u]$ for $t < u$, whenever u is the end of an FOMC meeting. The coefficients in the policy rule are computed with the parameter estimates $\hat{\boldsymbol{\gamma}}$ in table 1 for any $t < u$. The choice of $u - t$ depends on the frequency at which yield data are available. For example, with weekly data $u - t = 1/52$, we get

$$E_t[\theta_{t+(1/52)}] = 0.0036 + 0.87\theta_t + 0.10s_t + 7.51v_t + 0.0033z_t. \quad (19)$$

On the right-hand side, we have the last observation of the state variables $\boldsymbol{X}_t = \boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma})$ before the FOMC meeting.

From the *t*-statistics on the intensity parameters in table 1, we know that the slope coefficient on the variable z is estimated most precisely. The macro information in z is backed out mostly from the two-year yield, suggesting that the Fed reacts to a forecast of the economy over the next two years. To interpret the size of the coefficient in front of z, consider a one-standard-deviation shock to z. The standard deviation of z is 0.47, so that the target moves up by $0.0033 \times 0.47 = 16$ bp. The process z is autocorrelated, which leads to positive autocorrelated target changes, or policy inertia. The second most important variable is the target. The target coefficient of 0.87 induces persistence, or interest rate smoothing. From the *t*-statistics in table 1, we know that the spread s and the volatility v do not enter the policy rule significantly. Their estimated coefficients are small, given their standard deviations of 42 bp and below 1 bp, respectively. A one-standard-deviation shock to these variables shifts the target by fewer than 5 bp. These findings make sense economically. The Fed does not seem to care about random fluctuations of the short rate around the target or any heteroskedasticity in these fluctuations.

The coefficient estimates in (19) are roughly consistent with ordinary least squares (OLS) estimates. Unrestricted OLS runs the target at FOMC meetings on model-implied factors $\boldsymbol{X}_t = \boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma})$. The resulting

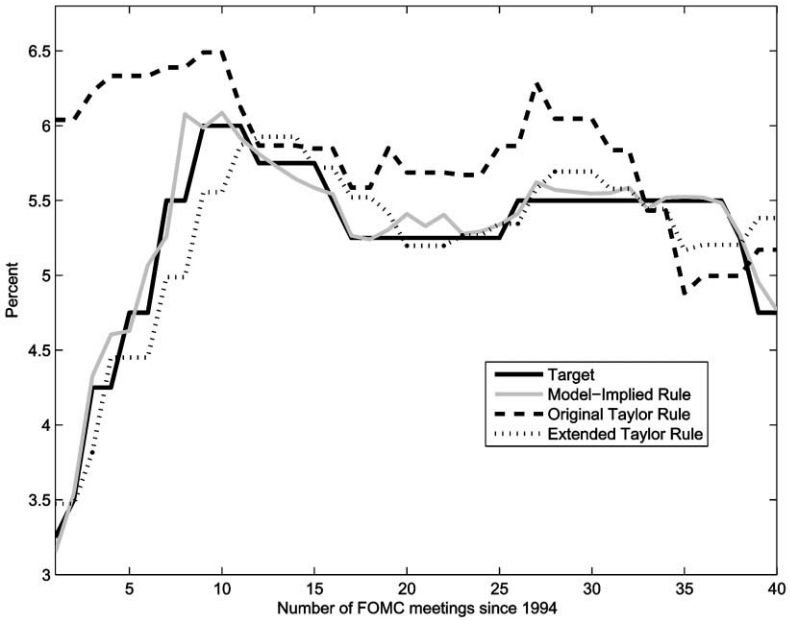


FIG. 3.—Target, model-implied policy rule, original Taylor rule, and extended Taylor rule at each of the 40 FOMC meetings (eight meetings per year) between 1994 and 1998.

intercept is 0.0015, and the slope coefficients are 0.89, 0.11, 9.04, and 0.0017.

While the factors are backed out from yield data, we cannot simply run the target on yields $\boldsymbol{Y}_t$, because the map $\boldsymbol{g}$ is nonlinear. But it turns out that not much is lost by ignoring this nonlinearity: the fitted values of a regression on $\boldsymbol{X}_t$ and a regression on $\boldsymbol{Y}_t$ differ maximally by 0.3 bp over the entire sample. To use the rule in practice, the Fed's staff can therefore run OLS with up-to-the-minute data on $\boldsymbol{Y}_t$. The rule thereby avoids Orphanides' (2001) critique of policy rules that are not based on real-time data (such as current GDP, which has yet to be released).

Figure 3 compares the policy rule (19) to Taylor rules. The left-hand-side variable in these rules is the quarterly averaged federal funds rate and not the target. But at this low frequency, the difference between these two rates is negligible. The Taylor rule uses two right-hand-side variables: inflation and the output gap. Inflation is measured as annual log changes in the GDP deflator, whereas the output gap is the percentage deviation of real GDP from its trend (based on a Hodrick-Prescott filter applied to quarterly data since 1947:1). The original Taylor rule is based on the coefficients proposed by Taylor (1993): 3 + 1.5 × inflation + 0.5 × gap. The estimated Taylor rule is based on esti-

mated OLS coefficients over the sample 1994–98. The extended Taylor rule adds the lagged federal funds rate to the right-hand side and also uses OLS to estimate the coefficients, following Clarida, Gali, and Gertler (2000). To mimic the decision process of the Fed, the graph plots the policy rule for each FOMC meeting given its value in the quarter in which the meeting took place, leaving us with 40 data points. The macro variables are taken from the current quarter, giving the Taylor-type rules the best chance at explaining the target movements (I tried various leads and lags).

By eyeballing, the model-implied rule seems to be a better description of the actual target. This is confirmed by the mean absolute difference between the actual target and the value of the target prescribed by the policy rule. For the Taylor rule, based on original and estimated coefficients, and the extended Taylor rule, the difference is 67, 43, and 22 bp, respectively. For the policy rule implied by the yield curve model, the difference is only 10 bp. Moreover, when we estimate the policy rule (19) using OLS, the difference is 9 bp, not much smaller.

From figure 3, we can see that the original Taylor rule does well as a general indication of Fed policy. For example, it was high in 1994 and low in 1998, even before the Fed moved the target. The same is true for the estimated Taylor rule (not included in the figure). In terms of mean absolute differences, the extended Taylor rule of course does better because it uses an additional right-hand-side variable. However, all Taylor rules lag behind, especially during times of repeated target moves in the same direction. This suggests that yields seem to be useful proxies for the information the Fed looks at.

D. Discrete Policy Choice Forecasts

Policy rules make continuous forecasts of target moves. To obtain discrete forecasts of whether the Fed will move the target or not at the next FOMC meeting, I also derive a discrete choice model using the estimated parameters. According to this model, the Fed randomizes over three possible policy choices at each FOMC meeting: up, down, or no move. Forecasting a particular choice means that the choice has the highest conditional probability. There have been only 40 FOMC meetings, so these forecasts suffer from small-sample noise. They provide, however, a device that helps to understand the model better. In particular, it is interesting to see whether the model tends to forecast moves in the wrong direction or whether the model tends to generate false positives by forecasting target moves when there is no move.

For each FOMC meeting, figure 4 plots the up and down probabilities conditional on information available right before the meeting. These probabilities are the empirical frequencies in 20,000 simulated samples.

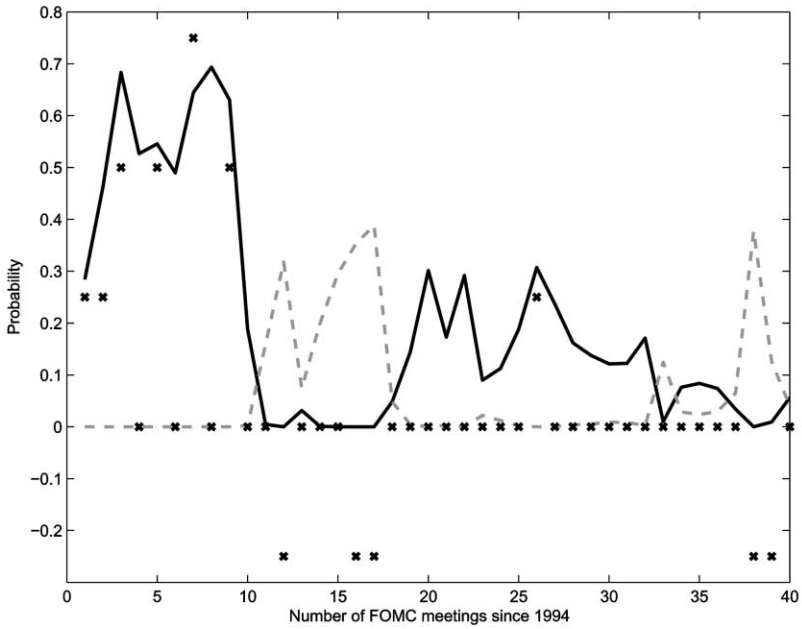


FIG. 4.—Conditional probabilities of up and down moves in the target at each of the 40 FOMC meetings (eight meetings per year) between 1994 and 1998. The solid line shows up moves, and the dashed line shows down moves. The *x*-marks indicate the actual target changes in percent.

The simulations start with the last observation on the factors $\boldsymbol{g}(\boldsymbol{Y}_t, \hat{\boldsymbol{\gamma}})$ before the FOMC meeting. Figure 4 shows that the conditional likelihood of moves up is very high at the end of 1994, when in fact the Fed increased the target in several steps, and again quite large around the target increase in March 1997. The conditional probability of moves down is high in 1995/96 and 1998, both years in which the Fed lowered the rate on several occasions.

Table 4 computes forecasts of choices from these probabilities. For example, the top-left number in the table means that there were four FOMC meetings for which the model forecasted an up move and the target really did go up. The bottom row shows that the forecasts from the model were correct for 30 out of 40 FOMC meetings. This is an overall correct forecasting percentage of 75 percent. The model never got the sign wrong. Each time the model forecasted up or down, the target either moved in that direction or did not move. Also, the model generated only two false positives. In other words, the model did not get 100 percent of the moves right because it tended to be "too cautious." It forecasted no move when there was a move, especially in the case of down moves.

TABLE 4
FORECASTING TARGET MOVES AT FOMC MEETINGS

	FORECAST				
ACTUAL	Up	No	Down	Correct	Total
Up	4	3	0	4	7
No	2	26	0	26	28
Down	0	5	0	0	5
Total	6	34	0	30	40

NOTE.—The sample goes from January 1994 to December 1998.

E. Yield Responses to Shocks

Figure 5 plots the yield coefficients $\mathbf{c}_X^y(t, T)$ from equation (13) as a function of maturity $T - t$. These yield coefficients can be interpreted as instantaneous responses of yields to the various shocks, because w^s, w^z, and w^v are orthogonal. The coefficients depend on calendar time, so I set t to the end of an FOMC meeting. This choice allows me to interpret the coefficient $c_\theta(t, T)$ on the target as the yield response to monetary policy shocks at FOMC meetings. From figure 5, this response is strong and falls with maturity only slowly. A one-percentage-point shock to the target shifts the one-month yield by 90 bp, the one-year yield by 60 bp, the two-year yield by 41 bp, and the five-year yield by 19 bp. These responses are roughly consistent with findings in Cochrane (1989), Evans and Marshall (1998), and Kuttner (2001). Long yields respond strongly to monetary policy shocks because shocks to the target die out only slowly under the risk-neutral measure. But eventually they do die out, so that long yields respond less than short ones. In the language of Litterman and Scheinkman (1991), the target θ is a "slope factor."

Figure 5 shows that the coefficient $c_z^y(t, T)$ (multiplied by 100 to make it comparable in size to c_s and c_θ) has a hump at two years. A one-standard-deviation shock to z shifts the one-month yield by 11 bp, the one-year yield by 38 bp, the two-year yield by 44 bp, and the five-year yield by 31 bp. The reason for the hump-shaped response of yields is the long half-life of shocks z: one year under the risk-neutral measure. A positive z-shock thus increases the risk-neutral probability of a rate hike not only at the next FOMC meeting but also at future FOMC meetings. Because of the anticipated cumulative effect of these hikes, intermediate yields respond more to z-shocks than short yields. At sufficiently long maturities, beyond two years, mean reversion in the target and the variable z causes shocks to have smaller and smaller impacts on longer and longer yields. The net effect is a hump-shaped coefficient on z with a peak at two years, which makes z a "curvature factor."

Figure 5 indicates that the response $c_s^y(t, T)$ of yields to spread shocks

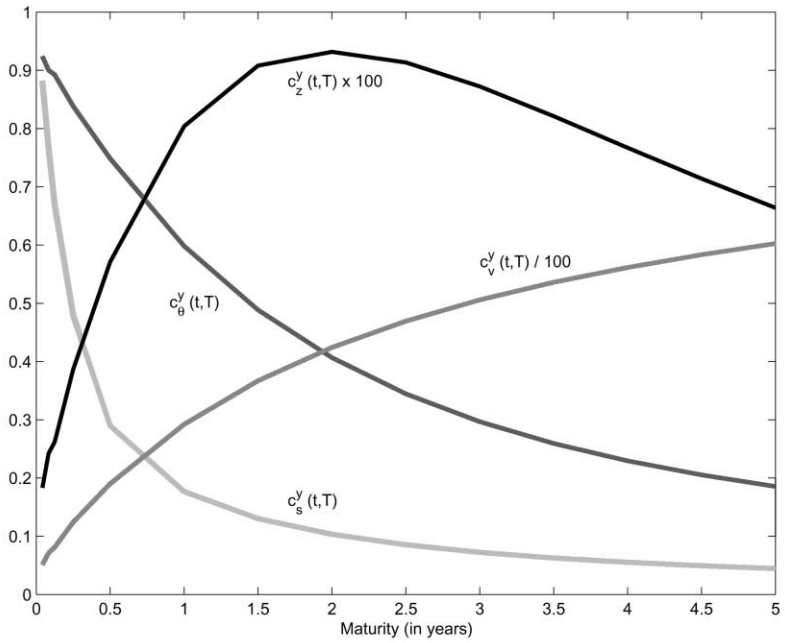


FIG. 5.—Responses of yields to monetary policy shocks $c^y_\theta(t, T)$, money market shocks $c^y_x(t, T)$, macroeconomic shocks $c^y_z(t, T)$, and volatility shocks $c^y_v(t, T)$. These coefficients are plotted as a function of maturity $T - t$, with t fixed to be the end of an FOMC meeting.

decreases very fast with maturity. A one-percentage-point shock to the spread shifts the one-month yield by 77 bp, the one-year yield by 18 bp, the two-year yield by 10 bp, and the five-year yield by 4 bp. In other words, both s and θ are "slope factors" but act on different parts of the yield curve, since the impact of money market noise dies off much faster with maturity (under the risk-neutral measure) than the impacts of monetary policy shocks.

Finally, figure 5 plots the coefficient $c^y_v(t, T)/100$, which is flat as a function of maturity. A one-standard-deviation shock to v shifts yields up by around 30–50 bp. The reason is that the persistence of the volatility factor v is extremely high under both measures. Shocks to volatility thus affect yields at all maturities. In this sense, v is a "level factor."

F. Snake-Shaped Volatility Curve

Figure 6 shows the volatility curve in the data, defined as the standard deviation of yield changes over the sample as a function of maturity. The curve connects the individual volatilities of various rates over a

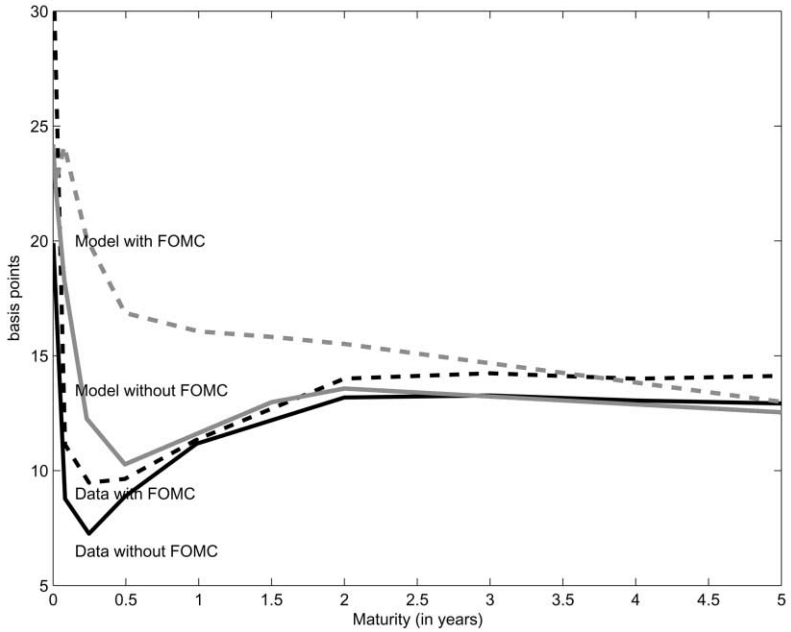


FIG. 6.—Snake shape and seasonality of the volatility curve. The four lines represent volatility curves during weeks with FOMC meetings and the remaining weeks computed from the data and the model as indicated.

weekly sample: Wednesday observations on the overnight repo rate and Thursday observations on the one-, three-, six-, and 12-month LIBOR rates and two-, three-, four-, and five-year swap rates. The curve is computed for weeks with an FOMC meeting and for the remaining weeks. The curve has a "snake shape": volatility is high at the very short end, rapidly decreases until maturities about three months, then increases until maturities of up to two years, and finally decreases again. The back of the snake, the hump at two years, has already been documented by Amin and Morton (1994).

Figure 6 also shows the volatility curve in simulated data from the estimated model. The 40,000 simulated samples are based on the actual FOMC meeting calendar. The simulated curve reproduces the overall snake-shaped pattern quite well. The model explains the back of the snake with *inertia in monetary policy*. Intuitively, volatility is due to two important types of shocks: macroeconomic shocks to z and money market noise. Macroeconomic shocks enter the model only through their impact on the conditional probability of a target move. These z-shocks increase the probability of a target move not only at the next FOMC meeting but also at subsequent meetings. Yields with medium maturities,

around two years, respond immediately to the anticipated cumulative effect of these pending target changes. We can see this from the hump in the yield coefficient $c_z^y(t, T)$ in figure 5. Money market noise is shocks to the spread between the overnight rate and the target. From figure 5, we know that these shocks generate reactions only in very short yields. The combined response of yields to these two types of shocks looks like a snake. These shocks are important for yields, so the snake shape carries over to the volatility curve.

Figure 6 suggests that volatility is higher during weeks with FOMC meetings, especially at the short end of the curve. The volatility in simulated data from the estimated model is also higher during those weeks. In fact, the simulated curve even overstates the seasonality somewhat for maturities around six months. The model explains the seasonality with *monetary policy shocks*. These are shocks to the target, which happen mostly at FOMC meetings. The seasonality is stronger for short yields because short yields respond more to monetary policy shocks than long yields. Again, we can see this from the downward-sloping coefficient $c_\theta^y(t, T)$ in figure 5. Monetary policy shocks dominate other shocks in weeks with FOMC meetings, which explains why the shape of the coefficient carries over to volatility.

VI. Conclusion

This paper shows that it helps to look at bond pricing and Fed policy jointly. The model formulates the target and yield dynamics in a consistent way. The estimation extracts information from both target and yield data. Target data improve the fit of the yield curve model and introduce important seasonalities around FOMC meetings. Data on long yields, especially yields with maturities around two years, enter the Fed's policy rule. There are many ways to go from here. An immediate extension is to investigate jumps and a more flexible volatility coefficient for the pricing kernel. Another extension is to include macroeconomic variables in the policy rule. These macro variables may capture information not contained in yields. Piazzesi (2001) takes first steps in this direction. Finally, the model can be applied to other central banks. For example, the European Central Bank and the Bank of England also announce their policy decisions at regularly scheduled meetings. All these extensions are left for future research.

Appendix A

Coefficients

To obtain the partial differential integral equation for bond prices, the following

notation is useful. The state vector $\boldsymbol{X}$ lives in $D \subset \mathbb{R}^4$ and solves the stochastic differential equation

$$d\boldsymbol{X}_t = \boldsymbol{\mu}_{\boldsymbol{X}}(\boldsymbol{X}_t)dt + \boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{X}_t)d\boldsymbol{w}_t + \boldsymbol{J}_{\boldsymbol{X}}(dN_t^U - dN_t^D), \tag{A1}$$

where $\boldsymbol{\mu}_{\boldsymbol{X}}: D \to \mathbb{R}^4$ is the drift, $\boldsymbol{\sigma}_{\boldsymbol{X}}: D \to \mathbb{R}^{4\times 4}$ is the volatility, and $\boldsymbol{J}_{\boldsymbol{X}} = [0.0025\ \boldsymbol{0}_{1\times 3}]^\top$ is the fixed jump size.

The bond price function F satisfies $F(\boldsymbol{X}, T, T) = 1$ at maturity, for all $\boldsymbol{X} \in D$. For $t \le T$, the function $F(\boldsymbol{X}, t, T)$ solves two different PDIEs, depending on whether t is within or outside an FOMC meeting interval. The difference between these PDIEs arises because of the jump intensities. During FOMC meetings, the intensities are (5). The resulting PDIE is

$$\begin{aligned} 0 = {} & F_t(\boldsymbol{X}, t, T) + F_{\boldsymbol{X}}(\boldsymbol{X}, t, T)[\boldsymbol{\mu}_{\boldsymbol{X}}(\boldsymbol{X}) - \boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{X})\boldsymbol{\sigma}_{\xi}(\boldsymbol{X})^\top] \\ & + \tfrac{1}{2}\mathrm{tr}[F_{\boldsymbol{X}\boldsymbol{X}}(\boldsymbol{X}, t, T)\boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{X})\boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{X})^\top] - [1_{1\times 2}\ \boldsymbol{0}_{1\times 2}]\boldsymbol{X} \\ & + [\overline{\lambda} - \boldsymbol{\lambda}_{\boldsymbol{X}}^\top(\boldsymbol{X} - \overline{\boldsymbol{X}})][F(\boldsymbol{X} + \boldsymbol{J}_{\boldsymbol{X}}, t, T) - F(\boldsymbol{X}, t, T)] \\ & + [\overline{\lambda} + \boldsymbol{\lambda}_{\boldsymbol{X}}^\top(\boldsymbol{X} - \overline{\boldsymbol{X}})][F(\boldsymbol{X} - \boldsymbol{J}_{\boldsymbol{X}}, t, T) - F(\boldsymbol{X}, t, T)], \end{aligned}$$

where tr denotes trace and F_t, $F_{\boldsymbol{X}}$, and $F_{\boldsymbol{X}\boldsymbol{X}}$ denote partial derivatives. Outside of FOMC meetings, the intensity of target moves is constant. Their values are set equal to their empirical frequency of one move per five years, or 0.2. The PDIE is then

$$\begin{aligned} 0 = {} & F_t(\boldsymbol{X}, t, T) + F_{\boldsymbol{X}}(\boldsymbol{X}, t, T)[\boldsymbol{\mu}_{\boldsymbol{X}}(\boldsymbol{X}) - \boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{X})\boldsymbol{\sigma}_{\xi}(\boldsymbol{X})^\top] \\ & + \tfrac{1}{2}\mathrm{tr}[F_{\boldsymbol{X}\boldsymbol{X}}(\boldsymbol{X}, t, T)\boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{X})\boldsymbol{\sigma}_{\boldsymbol{X}}(\boldsymbol{X})^\top] - [1_{1\times 2}\ \boldsymbol{0}_{1\times 2}]\boldsymbol{X} \\ & + 0.2[F(\boldsymbol{X} + \boldsymbol{J}_{\boldsymbol{X}}, t, T) - F(\boldsymbol{X}, t, T)] + 0.2[F(\boldsymbol{X} - \boldsymbol{J}_{\boldsymbol{X}}, t, T) - F(\boldsymbol{X}, t, T)]. \end{aligned}$$

Guess a solution of the form $F(\boldsymbol{X}, t, T) = \exp[\bar{c}(t, T) + \boldsymbol{c}_{\boldsymbol{X}}(t, T)^\top \boldsymbol{X}]$. The PDIEs must hold for all $\boldsymbol{X} \in D$, which I assumed contains an open set, so that I can apply the usual method of undetermined coefficients, which equates the coefficients of $\boldsymbol{X}$ and the constant terms to zero. The coefficients satisfy two systems of ordinary differential equations (ODEs). During subintervals with FOMC meetings, the ODEs are

$$\begin{aligned} \frac{d\bar{c}}{dt} = {} & -c_v\kappa_v\bar{v} + c_z q_z - 0.5c_z^2 + 2\overline{\lambda} - (\overline{\lambda} - \boldsymbol{\lambda}_{\boldsymbol{X}}^\top\overline{\boldsymbol{X}})\exp(0.0025c_\theta) \\ & - (\overline{\lambda} + \boldsymbol{\lambda}_{\boldsymbol{X}}^\top\overline{\boldsymbol{X}})\exp(-0.0025c_\theta), \\ \frac{dc_\theta}{dt} = {} & 1 - \lambda_\theta[\exp(0.0025c_\theta) - \exp(-0.0025c_\theta)], \\ \frac{dc_s}{dt} = {} & 1 + \kappa_s c_s - \lambda_s[\exp(0.0025c_\theta) - \exp(-0.0025c_\theta)], \\ \frac{dc_v}{dt} = {} & q_s c_s + (\kappa_v + q_v\sigma_v^2)c_v - 0.5c_s^2 - 0.5c_v^2\sigma_v^2 - \lambda_v[\exp(0.0025c_\theta) \\ & - \exp(-0.0025c_\theta)], \\ \frac{dc_z}{dt} = {} & \kappa_z c_z - \lambda_z[\exp(0.0025c_\theta) - \exp(-0.0025c_\theta)], \end{aligned}$$

where I suppress the dependence on t and T.

Outside of FOMC meeting intervals, the ODEs are

$$\frac{d\bar{c}}{dt} = -c_v\kappa_v\bar{v} + c_zq_z - 0.5c_z^2 + 2 \times 0.2 - 0.2\exp(0.0025c_\theta) - 0.2\exp(-0.0025c_\theta),$$

$$\frac{dc_\theta}{dt} = 1,$$

$$\frac{dc_s}{dt} = 1 + \kappa_sc_s,$$

$$\frac{dc_v}{dt} = q_sc_s + (\kappa_v + q_v\sigma_v^2)c_v - 0.5c_s^2 - 0.5c_v^2\sigma_v^2,$$

$$\frac{dc_z}{dt} = \kappa_zc_z.$$

The computation of $\bar{c}(t, T)$ and $\boldsymbol{c}_{\boldsymbol{X}}(t, T)$ is recursive. The algorithm divides the time between t and T into subintervals with and without FOMC meetings. The algorithm starts at T with terminal conditions $\bar{c}(T, T) = 0$ and $\boldsymbol{c}_{\boldsymbol{X}}(T, T) = 0$ and works its way backward up to the present $t \leq T$. Each time an FOMC meeting starts or ends, the current coefficients turn into terminal conditions, and the algorithm switches to the now-relevant ODEs. Some of these coefficients (such as $c_\theta(t, T)$ outside of meetings) can be easily solved analytically. Runge-Kutta methods solve the others numerically (in MATLAB, the relevant command is "ode45").

Appendix B

Simulations with Jumps

The state vector $\boldsymbol{X}$ contains the jump process θ. Starting from $\boldsymbol{x}$ at time τ, we can simulate $\boldsymbol{X}$ given by (A1) with the scheme

$$\Delta\hat{\boldsymbol{X}}^{\boldsymbol{x}}_{t+h} = \boldsymbol{\mu}_{\boldsymbol{X}}(\hat{\boldsymbol{X}}^{\boldsymbol{x}}_t)h + \sqrt{h}\boldsymbol{\sigma}_{\boldsymbol{X}}(\hat{\boldsymbol{X}}^{\boldsymbol{x}}_t)\boldsymbol{\epsilon}_{t+h} + \boldsymbol{J}_{\boldsymbol{X}}(\eta^U_{t+h} - \eta^D_{t+h}),$$

$$\hat{\boldsymbol{X}}^{\boldsymbol{x}}_\tau = \boldsymbol{x}, \tag{B1}$$

where $\boldsymbol{\epsilon}_{t+h}$ is independently and identically distributed standard normal, and η^U_{t+h} and η^D_{t+h} are independent Bernoulli variables with probabilities $\lambda^U_t h$ and $\lambda^D_t h$, respectively.

The simulations determine target changes by a "three-sided die." The three sides are "up" (U, meaning $\theta_{t+h} - \theta_t = 0.0025$), "down" ($D$, meaning $\theta_{t+h} - \theta_t = -0.0025$), and "no change" (0, meaning $\theta_{t+h} = \theta_t$). Their conditional probabilities at time $t + h$ are approximately

$$p^U_{t+h} = \lambda^U_t h(1 - \lambda^D_t h),$$

$$p^D_{t+h} = \lambda^D_t h(1 - \lambda^U_t h),$$

and

$$p^0_{t+h} = 1 - p^U_{t+h} - p^D_{t+h}.$$

In practice, I replace $\lambda^j_t h$ with $1 - \exp(-\lambda^j_t h)$ for $j = U, D$ to make sure that the probabilities behave well across all simulations.

The choice of h is important, especially with time-varying probabilities. At regular days, I set $h = 1/365$. At FOMC meetings, I need to further subdivide the day, because jump intensities can become large. For example, λ^U_t and λ^D_t take on values as high as 1,225 at the parameter $\hat{\boldsymbol{\gamma}}$ in table 1. At these values, a Bernoulli approximation that allows for only one jump during one FOMC meeting is not accurate. I therefore increase the number of Bernoulli trials during an FOMC meeting so that $h \geq (1/30)(1/365)$. To economize on the number of simulated steps (and thereby the computation time for the likelihood evaluation), I subdivide the FOMC meeting day into $H + 1$ intervals, where H is a multiple of five. For t during five subintervals of length

$$h = \left(\frac{H}{5}\right)\left(\frac{1}{H+1}\right)\left(\frac{1}{365}\right),$$

jumps are drawn from a Poisson distribution with constant parameter $\lambda^j_t h$ by truncating the distribution at $H/5$ jumps. In the last subinterval of length $h = 1/(H+1)$, a Bernoulli discretization is applied. I set $H = 30$, which is equivalent to 31 Bernoulli trials (with appropriately chosen success probability).

Appendix C

Simulated Maximum Likelihood with Jumps

The vector $\boldsymbol{X}^\theta$ denotes all variables in $\boldsymbol{X}$ other than the target θ, so that $\boldsymbol{X}_t = (\theta_t\ \boldsymbol{X}^\theta_t)^\top$. For the moment, suppose that the jump intensities are always "active," equal to (5), so that there is no time dependency introduced by FOMC meetings. The Monte Carlo approximation of the likelihood function in (17) is based on

$$\hat{f}_{\boldsymbol{X}}(\boldsymbol{X}_t,\ t|\boldsymbol{x},\ \tau) = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=\{U,D,0\}} \phi(\boldsymbol{X}^\theta_t,\ t|\theta_t,\ \hat{\boldsymbol{X}}^x_{t-h}[j],\ t-h)\hat{p}^k_t[j]1_{k,t}[j], \tag{C1}$$

where $\phi(\cdot, t|\hat{\boldsymbol{X}}^x_{t-h}[j],\ t-h)$ is the Gaussian density of $\boldsymbol{X}_t$ at time t conditional on the value $\hat{\boldsymbol{X}}^x_{t-h}[j]$ at time $t-h$; $\hat{\boldsymbol{X}}^x_{t-h}[j]$ denotes the jth simulated path from the scheme (B1); $1_{k,t}[j]$ is the indicator for the kth side of the die at time t in the jth simulation; and $\hat{p}^k_t[j]$ is based on $\hat{\boldsymbol{X}}^x_{t-h}[j]$. Let $\hat{\theta}^x_{t-h}$ be the target component of $\hat{\boldsymbol{X}}^x_{t-h}$. If the simulated target $\hat{\theta}^x_{t-h}$ at time $t-h$ cannot reach the observed time t value of the target in at most one jump, that simulation is assigned zero likelihood.

For the case of time-dependent intensities that are "activated" only during FOMC meeting day intervals $[t_{i-1},\ t_i]$, we can construct analogues of (C1) as follows. As long as the observation time t lies within a meeting day interval, in that $t_{i-1} \leq t < t_i$, the approximation (C1) itself still applies. If the observation time t is made outside an FOMC meeting, however, then we need to replace the Bernoulli density terms with an indicator function for sample paths leading up to the actual value of the target at t:

$$f_{\boldsymbol{X}}(\boldsymbol{X}_t,\ t|\boldsymbol{x},\ \tau) \approx \frac{1}{J}\sum_{j=1}^{J} \phi(\boldsymbol{X}^\theta_t,\ t|\hat{\boldsymbol{X}}^x_{t-h}[j],\ t-h)1_{\theta_t=\hat{\theta}^x_{t-h}[j]}. \tag{C2}$$

In (C2), jumps in the target enter the SML objective function only through the indicator function and the simulated values $\hat{\boldsymbol{X}}^{x}_{t_i-h}[j]$. This creates a serious problem when maximizing the objective: For a given (finite) number J of simulations, a small change in the parameter value does not necessarily affect the average number of jumps across simulations and may thus leave the value of the likelihood function unchanged. Only changes that are large enough to affect the number of simulated jumps change the objective function, but possibly by a large amount. In order to overcome this discontinuity, an alternative to (C2) is constructed as follows. The joint conditional density of factors can be written in the form

$$f_{\boldsymbol{X}}(\boldsymbol{X}_t,\ t|\boldsymbol{x},\ \tau) = f_\theta(\theta_t,\ t|\boldsymbol{x},\ \tau) f_{\boldsymbol{X}^\theta|\theta}(\boldsymbol{X}^\theta_t,\ t|\theta_t,\ \boldsymbol{x},\ \tau). \tag{C3}$$

The first term of equation (C3) can be approximated by

$$\begin{aligned} f_\theta(\theta_t,\ t|\boldsymbol{x},\ \tau) &\approx \frac{S}{J} \\ &\approx \frac{1}{J}\sum_{s=1}^{J} f_\theta(\theta_t,\ t|\,\hat{\boldsymbol{X}}^{x}_{t_i-h}[j],\ t_i - h) \\ &\approx \frac{1}{J}\sum_{j=1}^{J}\sum_{k=U,D,0} \hat{p}^k_{t_i}[j]1_{k,t}[j], \end{aligned}$$

where S denotes the total number of simulated paths that resulted in the observed value θ_t. In words, S/J is the frequency of "correctly simulated" target values in the simulations (starting with $\boldsymbol{x} = \boldsymbol{X}_\tau$), and the expression in the last row weighs the simulated paths by their likelihoods. Small changes in the parameters now affect the conditional probability $\hat{p}^k_{t_i}$, so that the likelihood is no longer discontinuous.

The second term in (C3) can be approximated by

$$\begin{aligned} f_{\boldsymbol{X}^\theta|\theta}(\boldsymbol{X}^\theta_t,\ t|\theta_t,\ \boldsymbol{x},\ \tau) &= \frac{f_{\boldsymbol{X}}(\boldsymbol{X}_t,\ t|\boldsymbol{x},\ \tau)}{f_\theta(\theta_t,\ t|\boldsymbol{x},\ \tau)} \\ &\approx \frac{1}{S}\sum_{j=1}^{J} \phi(\boldsymbol{X}^\theta_t,\ t|\,\hat{\boldsymbol{X}}^{x}_{t-h}[j],\ t-h)1_{\theta_t=\hat{\theta}^{x}_{t-h}[j]}. \end{aligned}$$

To evaluate the likelihood function, I simulate $J = 5{,}000$ paths of $\boldsymbol{X}$. The simulations use antithetic variates, which means that for each of the new pseudo-random Gaussian $\epsilon[j]$ and uniform $u[j]$, the antithetic variates $-\epsilon[j]$ and $1 - u[j]$ are used as a subsequent scenario. Like any simulation-based technique, SML is computationally intensive. The numerical optimization procedure is based on the Nelder-Mead simplex method, starting a gradient-based parameter search only after the simplex algorithm has collapsed.

Appendix D

Accuracy Check

The model does not impose a positivity constraint on intensities or the target. To check its approximation accuracy, I compute true zero-coupon yields $Y_0(t, T)$ with Monte Carlo integration. Starting at some value $\boldsymbol{x}$ for the factors at time

TABLE D1
APPROXIMATION ERRORS (in Basis Points)

	6-MONTH		2-YEAR		5-YEAR	
	$\hat{\boldsymbol{\gamma}}$ (1)	$\hat{\boldsymbol{\gamma}}_c$ (2)	$\hat{\boldsymbol{\gamma}}$ (3)	$\hat{\boldsymbol{\gamma}}_c$ (4)	$\hat{\boldsymbol{\gamma}}$ (5)	$\hat{\boldsymbol{\gamma}}_c$ (6)
Mean	2.1	2.6	1.7	2.2	1.9	1.5
Average standard error	.7	.6	1.1	1.1	1.8	1.6

NOTE.—The first row in this table presents mean absolute approximation errors $|\tilde{Y}_0(t, T) - Y_0(t, T)|$ in basis points over the weekly sample from January 1994 to December 1998 using parameter values $\hat{\boldsymbol{\gamma}}$ from table 1. The second row reports average standard errors of the Monte Carlo approximation of true yields $Y_0(t, T)$. They are obtained using the delta method by viewing the simulated bond price $\hat{P}(t, T)$ at time t as the estimated mean of an independently and identically distributed population of random variables $\exp(-\sum_{i=t}^{T-h} \hat{r}_i[j]h)$. The table reports the average standard errors over the sample.

t, the computation simulates J paths of the short rate $\hat{r}_i^x$ for times $i = t + h$, $t + 2h, \ldots, T - h$. The true yield is

$$Y_0(t, T) \doteq -\frac{\ln \hat{P}(t, T)}{T - t},$$

where

$$\hat{P}(t, T) = \frac{1}{J}\sum_{j=1}^{J} \exp\left(-\sum_{i=t}^{T-h} \hat{r}_i^x[j]h\right).$$

In the calculations, I set $J = 20{,}000$ and $h = 1/365$ and divide FOMC meeting days further into 30 intervals. Given these choices, the standard errors of the Monte Carlo approximation of the true yields for even the five-year yield (reported in the second row of table D1) are sufficiently small, from 0.6 to 1.8 bp.

At the same value $\boldsymbol{x}$, the model implies zero-coupon yields $\tilde{Y}_0(t, T)$ according to equation (13). The approximation errors $\tilde{Y}_0(t, T) - Y_0(t, T)$ are evaluated at typical $\boldsymbol{x}$-values, which are the factors $\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{Y}_t, \boldsymbol{\gamma})$ implied by the model at the estimated $\boldsymbol{\gamma}$. The FOMC calendar may introduce seasonalities into these errors. The first row in table 1 therefore reports average absolute approximation errors $|\tilde{Y}_0(t, T) - Y_0(t, T)|$ over the entire sample. For each maturity, columns 1, 3, and 5 are based on unconstrained estimates $\hat{\boldsymbol{\gamma}}$ and columns 2, 4, and 6 are based on constrained estimates $\hat{\boldsymbol{\gamma}}_c$ that solve (18). We can see that mean absolute errors are only around 1–3 bp. Also, the approximation errors using unconstrained estimates are not much different from those using constrained estimates.

References

Amin, Kaushik I., and Andrew Morton. 1994. "Implied Volatility Functions in Arbitrage-Free Term Structure Models." *J. Financial Econ.* 35 (April): 141–80.

Babbs, Simon H., and Nick J. Webber. 1993. "A Theory of the Term Structure with an Official Interest Rate." Manuscript, Cass Bus. School, London.

Bagliano, Fabio C., and Carlo A. Favero. 1998. "Measuring Monetary Policy with VAR Models: An Evaluation." *European Econ. Rev.* 42 (June): 1069–1112.

Balduzzi, Pierluigi, Giuseppe Bertola, and Silverio Foresi. 1997. "A Model of Target Changes and the Term Structure of Interest Rates." *J. Monetary Econ.* 39 (July): 223–49.

Bradsher, Keith. 1994. "Federal Reserve, Changing Course, Raises a Key Rate." *New York Times* (February 5), p. 1.

Campbell, John Y., and Robert J. Shiller. 1991. "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Rev. Econ. Studies* 58 (May): 495–514.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1999. "Monetary Policy Shocks: What Have We Learned and to What End?" In *Handbook of Macroeconomics*, vol. 1A, edited by John B. Taylor and Michael Woodford. Amsterdam: North-Holland.

Clarida, Richard, Jordi Gali, and Mark Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Q.J.E.* 115 (February): 147–80.

Cochrane, John H. 1989. "The Return of the Liquidity Effect: A Study of the Short-Run Relation between Money Growth and Interest Rates." *J. Bus. and Econ. Statis.* 7 (January): 75–83.

Dai, Qiang, and Kenneth J. Singleton. 2000. "Specification Analysis of Affine Term Structure Models." *J. Finance* 55 (October): 1943–78.

Das, Sanjiv R. 2002. "The Surprise Element: Jumps in Interest Rates." *J. Econometrics* 106 (January): 27–65.

Duffie, Darrell, and Rui Kan. 1996. "A Yield-Factor Model of Interest Rates." *Math. Finance* 6 (October): 379–406.

Duffie, Darrell, and Kenneth J. Singleton. 1997. "An Econometric Model of the Term Structure of Interest-Rate Swap Yields." *J. Finance* 52 (September): 1287–1321.

Eichenbaum, Martin, and Charles L. Evans. 1995. "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates." *Q.J.E.* 110 (November): 975–1009.

Evans, Charles L., and David A. Marshall. 1998. "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory." *Carnegie-Rochester Conf. Ser. Public Policy* 49 (December): 53–111.

Fama, Eugene F., and Robert R. Bliss. 1987. "The Information in Long-Maturity Forward Rates." *A.E.R.* 77 (September): 680–92.

Farnsworth, Heber, and Richard Bass. 2003. "The Term Structure with Semi-credible Targeting." *J. Finance* 58 (April): 839–65.

Hamilton, James D. 1996. "The Daily Market for Federal Funds." *J.P.E.* 104 (February): 26–56.

Hamilton, James D., and Oscar Jordà. 2002. "A Model of the Federal Funds Rate Target." *J.P.E.* 110 (October): 1135–67.

Johannes, Michael. 2004. "The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models." *J. Finance* 59 (February): 227–60.

Kuttner, Kenneth N. 2001. "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market." *J. Monetary Econ.* 47 (June): 523–44.

Litterman, Robert, and José A. Scheinkman. 1991. "Common Factors Affecting Bond Returns." *J. Fixed Income* 1 (June): 54–61.

Longstaff, Francis A., and Eduardo S. Schwartz. 1992. "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model." *J. Finance* 47 (September): 1259–82.

Meulendyke, Ann-Marie. 1998. *U.S. Monetary Policy and Financial Markets*. New York: Fed. Reserve Bank.

Meyer, Laurence H. 1998. "Come with Me to the FOMC." Gillis Lecture (April 2), Willamette Univ., Salem, OR. http://www.federalreserve.gov/boarddocs/speeches/1998/199804022.htm.

Orphanides, Athanasios. 2001. "Monetary Policy Rules Based on Real-Time Data." *A.E.R.* 91 (September): 964–85.

Pedersen, Asger R. 1995. "A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations." *Scandinavian J. Statis.* 22 (March): 55–71.

Piazzesi, Monika. 2001. "An Econometric Model of the Yield Curve with Macroeconomic Jump Effects." Working Paper no. 8246 (April), NBER, Cambridge, MA.

———. 2004. "Affine Term Structure Models." Manuscript, Grad. School Bus., Univ. Chicago.

Rudebusch, Glenn D. 1995. "Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure." *J. Monetary Econ.* 35 (April): 245–74. [Erratum 36 (December 1995): 679.]

Santa-Clara, Pedro. 1995. "Simulated Likelihood Estimation of Diffusions with an Application to the Short Term Interest Rate." PhD diss., Insead, Paris.

Taylor, John B. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conf. Ser. Public Policy* 39 (December): 195–214.