An Econometric Model of the Yield Curve with Macroeconomic Jump Effects

Monika Piazzesi*
UCLA and NBER

First Draft: October 15, 1999
This Draft: April 10, 2001

Abstract

This paper develops an arbitrage-free time-series model of yields in continuous time that incorporates central bank policy. Policy-related events, such as FOMC meetings and releases of macroeconomic news the Fed cares about, are modeled as jumps. The model introduces a class of linear-quadratic jump-diffusions as state variables, which allows for a wide variety of jump types but still leads to tractable solutions for bond prices. I estimate a version of this model with U.S. interest rates, the Federal Reserve's target rate, and key macroeconomic aggregates. The estimated model improves bond pricing, especially at short maturities. The "snake-shape" of the volatility curve is linked to monetary policy inertia. A new monetary policy shock series is obtained by assuming that the Fed reacts to information available right before the FOMC meeting. According to the estimated policy rule, the Fed is mainly reacting to information contained in the yield-curve. Surprises in analyst forecasts turn out to be merely temporary components of macro variables, so that the "hump-shaped" yield response to these surprises is not consistent with a Taylor-type policy rule.

*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, Thomas 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 seminar participants at Carnegie-Mellon, Chicago, Columbia, Cornell, Harvard, MIT, the NBER Asset Pricing Meeting 2000, the NBER Summer Institute 2000, Northwestern, NYU, Princeton, Rochester, Stanford, Tel Aviv, Tilburg, Toulouse, UBC, UCL, UCLA, USC, the WFA 2000, the Workshop on Mathematical Finance at Stanford 2000, and Yale. The financial support of Doctoral Fellowships from the Bradley and Alfred P. Sloan Foundations are gratefully acknowledged. The latest version can be downloaded from www.personal.anderson.ucla.edu/monika.piazzesi or requested from piazzesi@ucla.edu.

1 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, most prominently nonfarm payroll employment.[1] The FOMC is well aware of being closely watched by the markets, and extracts information about the current state of the economy from the current yield curve. This yield-based information may underlie the FOMC's policy decisions.

These observations suggests that models of the yield curve should take into account monetary policy actions by the Fed. The extensive term-structure literature in finance, however, builds models around a few latent state variables which are backed out from yield data. This statistical description of yields only offers 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 property especially applies to the very short end of the yield curve, as most studies avoid dealing with the extreme volatility and the large outliers at certain calendar days of short-rate data (as documented by Hamilton (1996)).

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 (see Christiano, Eichenbaum and Evans (2000) for a survey), 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' - this seems to exclude the first question. And financial markets immediately react to FOMC announcements - which seems to exclude the second.[2]

This paper attempts to kill these two birds with one stone. The idea is that at high data-sampling frequencies, information about the exact timing of policy-related events, such as FOMC meetings and macroeconomic releases, can be used to improve bond pricing and to identify monetary policy shocks. I therefore construct a continuous-time model of the joint distribution of bond yields, the interest-rate target set by the Fed and macroeconomic

[1] For Garch-type models of volatility with release day dummies and macroeconomic news surprises, see Jones, Lamont, and Lumsdaine (1996) and Li and Engle (1998). For regressions of yield changes on news surprises, see Balduzzi, Elton, and Green (1998) and Fleming and Remolona (1997).

[2] Such a recursive identification assumption is made by Eichenbaum and Evans (1995) to study the impact of monetary policy shocks on quarterly exchange rates. Eichenbaum and Evans opt for the first of the two possible assumptions, namely that the Fed does not react to current exchange rates.

variables the Fed cares about. The model imposes the cross-equation restrictions implied by the absence of arbitrage and respects the timing of policy moves at FOMC meetings and releases of macroeconomic variables that are likely to affect future Fed policy. These policy-related events happen at distinct points in time and are therefore captured as jumps. Although these jumps can be irregularly spaced and depend on the state of the economy, the model has closed-form solutions for bond prices. This feature allows an estimation with long-maturity yield data, which is particularly important in the context of monetary policy.[3] I estimate various versions of this model with data since 1994 on LIBOR (London Interbank Offered Rate) and swap rates in which the target rate and macroeconomic aggregates are observable factors, along with more traditional latent factors. The estimation is by the method of simulated maximum likelihood (Pedersen (1995), Brandt and Santa-Clara (2001)), extended here to the case of jumps.

There are four main estimation results. First, the model considerably improves the performance of existing yield-curve models with three latent state variables (such as Dai and Singleton (2000)), especially at the short end of the yield curve. Intuitively, the target set by the Fed provides a 'clean measure' of the short rate and therefore offers a tractable way to fix the short end of the yield-curve at a good position. This improvement can be achieved with a 4-factor model that includes (*i*) the target rate, (*ii*) a spread factor that measures deviations of the short rate from the target, (*iii*) a traditional stochastic volatility factor and (*iv*) a policy inertia factor. The short rate reverts quickly and continually to the target, suggesting that the spread factor captures money-market noise which does not matter for longer term yields. The target adjusts slowly toward the Fed's new desired target *only through jumps occurring at FOMC meeting days.* The likelihood of target-rate moves at FOMC meetings depends crucially on two factors: the current target and the inertia factor. The inertia factor, which is closely related to the 2-year yield, thus captures latent information contained in the yield curve that the Fed takes into account when conducting policy. Persistence in the target holds the target near its old value (*interest-rate smoothing*), thereby introducing *positive autocorrelation in the target-rate level.* The inertia factor slowly pulls the target toward the new desired value of the target (*policy inertia*). Shocks to the policy-inertia factor increase the likelihood of a target move, not only at the next meeting, but also at subsequent meetings. This leads to *positive autocorrelation in target-rate changes.*

Second, the paper documents the *snake-shape of the volatility curve* that can be seen in Figure 4, the standard deviation of yield changes as a function of maturity, and links this stylized fact to policy inertia. The cross-sectional response of yields to shocks in the

[3]Unrestricted regressions of short-rate changes on changes in policy rates have small sample problems because the Fed only moves the target infrequently. Moreover, they are haunted by reserve-maintenance-period effects, as most policy moves before 1994 happened on Thursdays after "settlement Wednesdays." After 1994, almost all policy moves were made at FOMC meetings, which are usually scheduled for either Tuesday or Wednesday. By placing the target in a yield-curve model, data on longer yields (which are little affected by settlement Wednesdays) provide additional information about the parameters governing the reaction of yields to target changes.

inertia factor has a hump at maturities around 2 years, as the anticipated cumulative effect of pending target changes is largest for those maturities. The combined response to money-market shocks and inertia-factor shocks has a snake-shaped pattern: high for very short maturities, rapidly decreasing until maturities of around 6 months, then increasing until maturities of up to 2 years, and finally decreasing again. As these shocks are important for yields, this snake-shaped pattern carries over to the volatility curve.[4] Shocks to the target rate in the post-1994 environment happen mostly at FOMC meetings and thereby introduce a seasonality into the volatility of yields.

Third, I obtain a new policy-shock series and policy-rule estimate by assuming that the Fed reacts to information known right before the FOMC meeting. This makes sense, because monetary policy is the most important, if not only, source of shocks to macroeconomic and yield variables right at the time of the FOMC announcement. This assumption identifies a *high-frequency policy rule* which is given by the expected value of the target rate at the FOMC announcement conditional on information before the meeting. While this information includes macroeconomic variables and yields of different maturities, the estimated rule shows that yields are more important for the Fed. As a description of target dynamics, the rule performs better than several benchmarks, including estimated versions of the Taylor rule (Taylor (1993)). This improvement in fit is not a simple consequence of adding explanatory variables to the rule, because of the cross-equation restrictions on the parameters. The extent of the improvement is due to the fact that the yield curve, especially information contained in yields with maturity around two years, reflects changes in the state of economy that the Fed is reacting to. This makes yields look like they are 'anticipating' target moves which results in highly significant coefficients of yields in the policy rule. The cross-sectional response of yields of different maturities to monetary policy shocks is monotonically decreasing in maturity.

Finally, macroeconomic release surprises turn out to be temporary components in macroeconomic variables, in the sense that the impulse response of macro variables to these shocks dies off after one month. Release surprises are here defined as difference between actual released values (nonfarm payroll employment and CPI inflation) and analyst forecasts. In a model in which the Fed reacts to current macroeconomic variables, this means that release surprises can affect the conditional probability of target moves at only those FOMC meetings that are scheduled before the next macro release. In other words, *release surprises are not innovations to inertia-type state variables.* In order to replicate the hump-shaped cross-sectional response of yields to release surprises, the propagation of these surprises would need to 'live longer.' This may be achieved by allowing for correlation between the release surprises and the policy-inertia factor.[5] Here, macroeconomic news have an impact on the

[4]The volatility hump at 2 years in the data is well known, see Litterman, Scheinkman and Weiss (1988).

[5]This amounts to attributing release surprises to latent variables, which avoids the need to specify the dynamics of macroeconomic variables and their connection to the yield curve. This strategy is followed in Fleming and Remolona (1999) who investigate the behavior of high-frequency yield data around macroeconomic news releases using a discrete-time Gaussian yield-curve model.

new desired target, but the FOMC only gradually implements this desired target over a number of meetings.

Yield models with low-dimensional state vectors which are backed out from yields go back to Merton (1974), Vasicek (1977) and Cox, Ingersoll, and Ross (1985, CIR). The theoretical frameworks of Duffie and Kan (1996) and El Karoui, Myneni and Viswanathan (1993) nest most of the specifications estimated today, as they provide tractable bond-pricing formulas. For empirical applications, see, for example, Anderson and Lund (1996), Dai and Singleton (2000) and Ahn, Dittmar, and Gallant (2001). Even in models in which the state variables are observable, they are usually not used in the estimation. For example, Pearson and Sun (1994) estimate a nominal version of the CIR model, in which inflation is specified as a factor, with yield data only. Campbell and Viceira (2000) is an exception, as they filter latent variables using yields and CPI inflation.

A small number of papers address the yield curve and monetary policy. Babbs and Webber (1993, 1996) and Farnsworth and Bass (2000) present theoretical yield-curve models that capture some aspects of monetary policy. These models do not have tractable solutions for yields and are therefore not taken to the data. Another group of papers estimates short-rate processes that reflect features of policy (Rudebusch (1995), Balduzzi, Das and Foresi (1996), Balduzzi, Bertola, Foresi and Klapper (1998), Honoré (1997), Hamilton and Jorda (2000), Das (2001)). The first three papers in this group use their estimated short-rate process to compute long yields by invoking the expectation hypothesis. These long yields are then taken to study the impact of Fed policy on the regression results by Campbell and Shiller (1991). The same is done by Konstantinov (2000) who estimates a regime switching model for the short rate with yields up to 1 year maturity. These papers do not estimate a fully fledged yield-curve model nor do they study monetary policy shocks.

While VARs also capture bond yields and macroeconomic variables, they give rise to a joint time-series process that implies arbitrage opportunities. Evans and Marshall (1998), for example, estimate various VARs that amount to an unrestricted regression of bond yields on different traditional monetary policy shock series. The assumption used to identify these shocks is that the Fed ignores both current and past yield-curve information. Another strand of literature uses a general equilibrium setting, which does imply the absence of arbitrage (for example, Pennacchi (1991), Berardi (1998) and Jiltsov and Buraschi (1999)). The approach I take also imposes no-arbitrage, while not requiring the specification and estimation of a structural model of the economy, which in any case would be problematic given the current state of empirical GE models of asset prices (see, for example, Hansen and Jagannathan (1991). For a recent survey, see Cochrane (1998)).

The paper is structured as follows. Section 2 provides some institutional background on the operating procedures of U.S. monetary policy. Section 3 presents the general theoretical framework with policy and the different versions of the model that are estimated. Section

4 describes the computation of the pricing formula, the simulation-based estimation technique, and the data. Sections 5 and 6 present estimation results obtained without and with macroeconomic variables, respectively.

2 The Timing and Size of Target Changes

Important changes in 1994 to Fed-policy operating procedures underlie the choice of sample period in this paper, which focuses on the policy framework in place today. The Fed conducts monetary policy by targeting 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 shows the fed funds market rate, the target rate together with LIBOR and swap rates from 1994 to 1998. The Figure illustrates that deviations of the fed funds rate from the target come in the form of short-lived spikes. (Section 4.5 provides a description of the target data that is used in this paper.) These spikes are usually associated with "settlement Wednesdays" and other special calendar effects, such as the end of the year.[6] Figure 1 also shows that target-rate changes are often followed by additional changes in the same direction. This feature will be referred to as "policy inertia".[7]

Starting with the first FOMC meeting of 1994, the Fed changed the timing and the size of target moves. This change 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-rate change and the preceding FOMC meeting. If, in a 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 re-targeting mainly at FOMC meeting days, with two exceptions highlighted in Figure 2. The first exception occurred on April 18, 1994 after high car sales in March, a leading business-cycle indicator. The financial press speculated that the surprise move was intended as a manifestation of authority by Alan Greenspan, as no vote was held on the move.[8] The second exception was decided upon in a conference call on October 15, 1998, and came in response to the Russian financial crises.[9] The lower row of graphs in Figure 2 shows the histogram of target changes

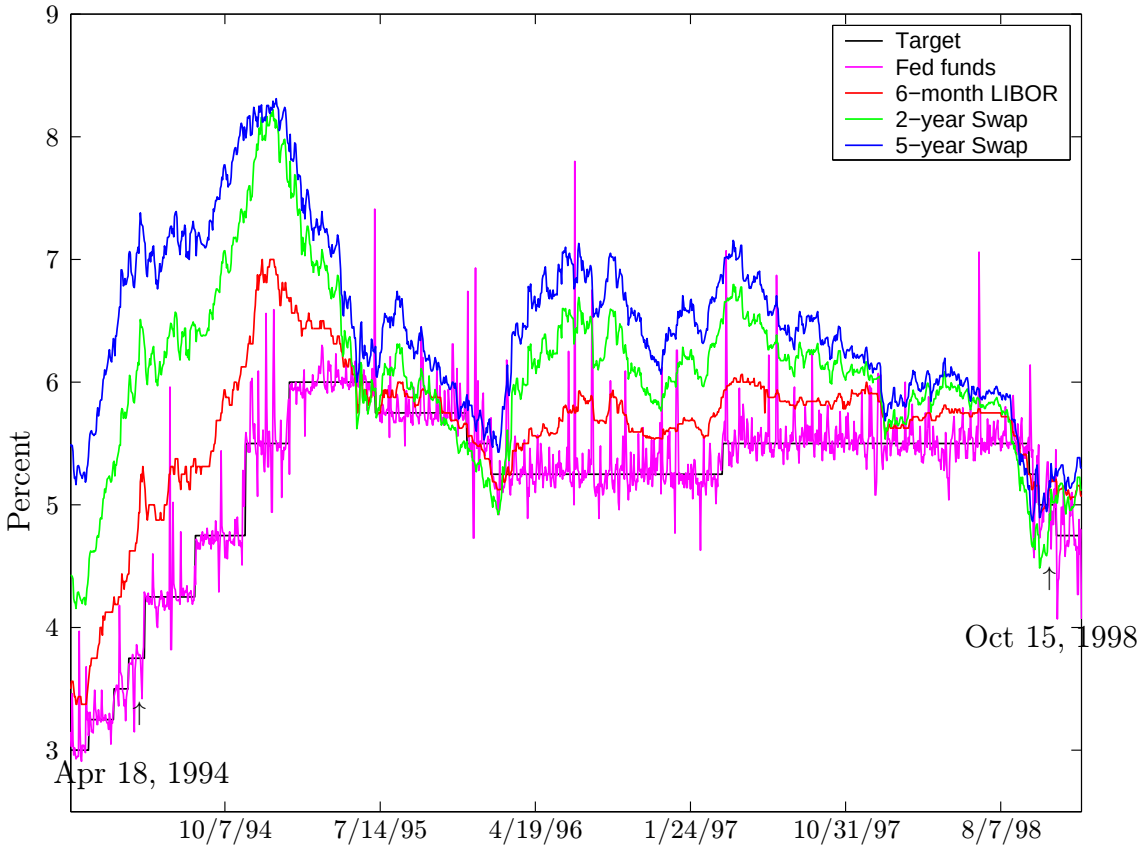
[6] During bi-weekly reserve-maintenance periods, banks must hold "good funds" in the form of cash or in accounts at the Fed, or be penalized. These reserve-maintenance periods end on "settlement Wednesdays".

[7] Policy inertia, as defined here, does not only apply to the short period considered in the sample. For example, from June 1989 to September 1992 the Fed lowered the target 25 times without a ever increasing it. Policy inertia may indicate the Fed's unwillingness to move the target immediately and entirely to its desired rate. Instead, the Fed adjusts the target in small steps to avoid possible policy mistakes, because of political motives, parameter uncertainty (Sack(1998)) or to affect long-maturity yields with minimal changes in short yields (Woodford (1999)).

[8] *The Financial Times*, April 19, 1994, page 3, "Greenspan plays an early hand: US rates rise" by Michael Prowse and *The New York Times*, April 19, 1994, page 1, "Fed again raises short-term rate on loans" by Keith Bradsher.

[9] *The New York Times*, October 16, 1998, page 1, "Federal Reserve cuts rates again; Wall St. surges" by

Figure 1: Daily Target, Fed funds, LIBOR and Swap Rates from 1994 to 1998



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.

3 Yield-Curve Model with Fed targeting

3.1 Dynamics of the Target

The target looks like a step function with steps that are multiples of 25 basis points. In a continuous-time model, these features call for a pure jump-process specification in which the target θ solves the stochastic differential equation (SDE)

$$d\theta(t) = 0.0025\ (dN_U - dN_D)\,, \tag{1}$$

Richard W. Stevenson.

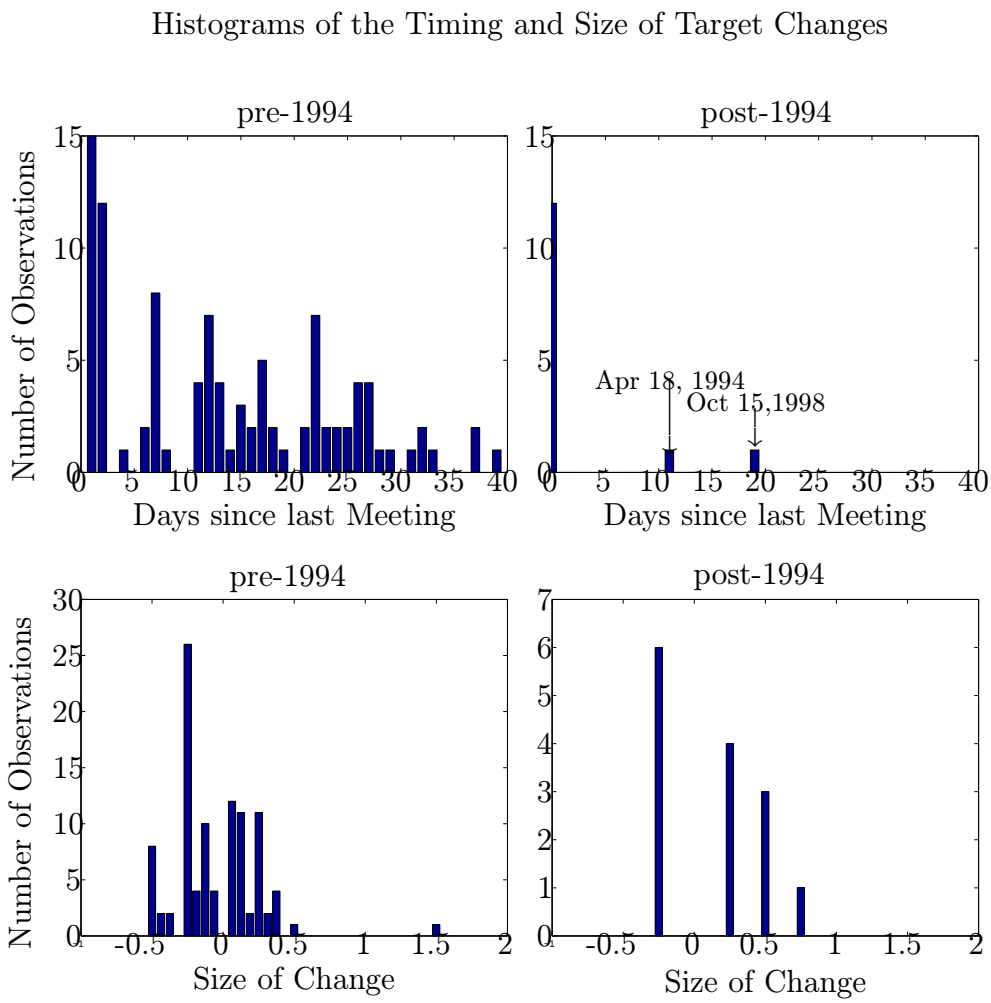


Figure 2: The first row of graphs show the histogram of days since the last FOMC meeting, for any given target rate change between 1984-1993 and 1994-1998. In the first subperiod, there have been a total of 100 target moves, while there were 14 in the second subperiod. The second row of graphs show the histogram of the size of target changes between 1984-1993 and 1994-1998.

where N_U and N_D are counting processes with stochastic intensities λ_U and λ_D, respectively. Jumps in N_U and N_D increase the target by 25 bp, while jumps in N_D lower the target by the same amount. Heuristically, the probability of a Poisson jump in N_U during the interval $[t, t+\Delta]$ conditional on information up to time t is given by $\lambda_U(t)\Delta$, and the expression for jumps in N_D is analogous. The econometrician has discrete observations on the difference between N_U and N_D, which means that he observes target moves that are multiples of 25 bp.

The Fed sets the target in response to the state of the economy X that lives in some state space $D \subset \mathbb{R}^N$. More precisely, the Fed's policy decision at time t is based on the value of X 'just before' time t, which is denoted $X(t-) = \lim_{s\uparrow t} X(s)$. The conditional probability of a target move varies according to the FOMC meeting calendar. The arrival rate of 'regular' target moves at FOMC meetings depends on the state X of the economy. Then there is a small and constant probability of an emergency move outside of FOMC meetings ('Peso events'). More precisely, FOMC meetings are taken to be intervals, the i-th meeting being $[\widetilde{t}_M(i), t_M(i)]$. The arrival rates λ_U and λ_D are given by

$$\lambda_j(t) = \begin{cases} \lambda_0^j + \lambda_X^j \cdot X(t-) & \text{for} \quad t \in [\tilde{t}_M(i), t_M(i)], \\ \lambda_P^j, & \text{otherwise}, \end{cases} \tag{2}$$

for $j = U$ ("up") and D ("down").

The analogy between the target process (1) and descriptions of Fed behavior in the Taylor rule literature can be seen more easily by defining the martingale $M = M_U - M_D$, where M_j is the compensated process given by $\{M_j(t) = N_j(t) - \int_0^t \lambda_j(t)dt \; ; t \geq 0\}$ for $j = U, D$ (see Brémaud (1981) for further details). Then the dynamics of the target can be rewritten as

$$d\theta(t) = \kappa(t)\left(\overline{\theta}(t-) - \theta(t-)\right) dt + 0.0025\, dM(t), \tag{3}$$

where the process $\overline{\theta}$ is a linear function of the state X with coefficients dictated by the intensity parameters. This representation shows that, during times at which the deterministic scalar $\kappa(t)$ is positive, the target is expected to move back towards $\overline{\theta}(t-)$. The expected value of $\theta(t)$ (and of the average target over some period) conditional on information at time $s \leq t$ is linear in $X(s)$. This expected value is thus directly comparable to the Taylor rule, for example, which links the average federal funds rate to measures of inflation and output at a quarterly frequency.[10]

To economize on parameters, the slope parameters are taken to be symmetric, in that $\lambda_X := \lambda_X^U = -\lambda_X^D$. Mean-reversion in the target is also imposed by assuming $\lambda_0^D = \lambda_0^U + 2\lambda_X \overline{x}$, where $\overline{x}$ denotes the long-run mean of X. The arrival rates of Peso events are fixed to their empirical frequency. There has been one up and one down move outside of FOMC meetings the 5 years from 1994 to 1998, so that I set $\lambda_P^U = \lambda_P^D = 0.2$. For given long-run mean parameters, there are thus $N+1$ free parameters: λ_0^U and λ_X.

[10] At a quarterly frequency, the difference between the average fed funds and average target rate is tiny. Given the convention that $[t, t+1]$ spans one year, the conditional expected value of the average target over $[s, t]$ can be computed exactly using $s - t = 0.25$.

3.2 General State Dynamics (including Target)

The state $X(t)$ at time t includes the target rate $\theta(t)$, some macro variables $m(t)$, and analyst forecasts $m_F(t)$ of these macro variables. All these state variables are observable. The remaining state variables are unobservable (or latent). Among these is the short rate and thus the spread $s(t) = r(t) - \theta(t)$. The main reason for this approach is that the spikes in the fed funds rate (see Figure 1) call for a detailed specification of the underlying seasonality similar to Hamilton (1996), which would add a large number of parameters. This complexity can be avoided by backing out the short rate from longer term yields.

The general dynamics of X are described by an SDE of the form

$$dX(t) = \mu(X(t))\,dt + \sigma(X(t))\,dW(t) + dJ(t), \tag{4}$$

where W is a vector of Brownian motions, μ is the drift, σ is the volatility, and J is a pure jump process. The drift $\mu(x)$ and the variance-covariance term $\sigma(x)\sigma(x)^\top$ are linear functions of the state x. Without jumps and with constant σ, this system may be thought of as a Gaussian vector-autoregression with mean rate of change μ. Taking the variance-covariance $\sigma\sigma^\top$ to be state-dependent allows for conditional heteroscedasticity. More importantly, the jumps J can capture 'macroeconomic jump effects', such as policy events and macroeconomic releases. Here, jumps do not necessarily capture 'large moves' in yields (as in the short-rate studies of Das (2001) and Johannes (2001)), but observable changes in X that happen at distinct points in time.

The specification of jumps differs from existing parametric models (Duffie and Kan (1996), El Karoui, Myneni and Viswanathan (1993)) to capture different operating procedures of monetary policy and macroeconomic news releases. The linearity assumption on μ and $\sigma\sigma^\intercal$ combined with parametric assumptions on the distribution of J defines a new class of continuous-time Markov processes labelled *linear-quadratic jump-diffusions (LQJDs).* To describe the properties of this class of processes, I partition the state vector as $X = (X_1, X_2)$, so that X_2 is a k_2-dimensional process, with $k_1 + k_2 = N$. Without loss of generality (k_2 may be zero), the subvector X_2 is assumed to be Gauss-Markov, which means that the lower k_2 components of $\Delta X(t)$ are always zero. In what follows, whenever the term 'quadratic' is used, it refers to squaring Gaussian processes (for details, see Appendix A). There can be jumps at

- *scheduled announcements* (deterministic points in time): The distribution of the jump size $\Delta X(t)$ conditional on information 'just before' the deterministic jump time t may depend on the state $X(t-)$. In particular, the moment generating function of the conditional distribution is an exponential linear-quadratic (LQ) function of $X(t-)$.

- *unscheduled announcements* (random points in time): Random jump times τ arrive with intensities which are LQ functions of the state. The distribution of $\Delta X(\tau)$ is state independent.

There is thus a dichotomy between state dependence of the jump-size distribution on the one hand and of jump timing on the other. This separation turns out to be crucial for bond prices to have closed form solutions. In this paper, I use scheduled announcements to model releases of employment and CPI numbers. The requirement on the moment generating function of these jumps is satisfied, for example, by a Gaussian distribution with LQ mean in $X(t-)$ and constant variance. The target process is based on jumps at unscheduled announcements with time-varying arrival intensities, because there is always a small probability of a 'Peso move' outside of FOMC meetings. Moreover, target changes come in discrete amounts and, even more importantly, a zero move always has positive probability. The distribution of target changes therefore does not satisfy the condition on the moment generating function of jumps at scheduled announcements.

3.3 Negative Probabilities of Target Moves?

Negative correlation between the conditional probabilities of up and down moves in the target is important for a description of the countercyclical interest-rate targeting by the Fed. The linear arrival intensities λ_U and λ_D in (2) allow for such negative correlation. While quadratic intensities also admit this feature, the linear specification has the further advantages that the map from states to yields is invertible and, even more importantly, that the target rate can revert to a mean. With a state-independent jump size distribution, the only way to achieve this is to allow the arrival rates to depend on θ. Since the target is not Gaussian and LQJDs only square Gaussian processes (as laid out in Appendix A), this dependence has to be linear. Invertibility of the pricing map means that the model can be estimated with a method that relies on the likelihood function of the state vector.

The main disadvantage of (2) is that intensities may become negative.[11] For the model to be well defined, I therefore assume that (2) defines the intensities of an *approximating yield-curve model*, while the true intensities are truncated versions given by $\max\{\lambda_j, 0\}, j = U, D$. These truncated linear intensities are outside of the LQJD class, and a closed-form solution for bond yields is no longer available. The quality of the approximating model can be checked at a given parameter using Monte-Carlo methods (see Section 5.1). The idea of using a tractable yield-curve model as an approximation usually also underlies papers that estimate nominal short-rate processes that may become negative (which are most papers in the literature, including this one). The true short rate in this case is given by $\max\{r, 0\}$.

[11] Negative correlation of intensities with positive domain force a violation of assumption Condition A in Duffie and Kan (1996). In the absence of jumps, the condition is sufficient for the existence of a solution to the stochastic differential equation (4) describing the state. As already noted by Duffie and Liu (2001), it is possible to square two Gaussian processes so that each variable takes only positive values, while allowing for arbitrary correlation. This idea is applied here to the case of jumps.

3.4 Other State Variables

All estimated setups share the feature that that the state vector X includes the target as observable variable and the spread $s = r - \theta$ as latent, with s mean-reverting to zero. The fat tails of the spread distribution can be captured with a stochastic volatility factor v. The most interesting variable is called *policy inertia factor* z, which proxies variables (in addition to s, θ and v) to which the Fed reacts when setting the target. The arrival intensities (2) of policy moves during FOMC meetings may depend on all these variables in X. The dynamics of these additional state variable solve

$$\begin{aligned} ds(t) &= -\kappa_s s(t)\, dt + \sqrt{v(t)} dW_s(t), \\ dv(t) &= \kappa_v(\overline{v} - v(t))\, dt + \sigma_v \sqrt{v(t)} dW_v(t), \\ dz(t) &= -\kappa_z z(t)\, dt + dW_z(t), \end{aligned} \tag{5}$$

where W_s, W_v and W_z are independent standard Brownian motions.

The four-dimensional state vector $X = (\theta, s, v, z)$ does not yet include macro variables. Section 6 tries link the latent information in z to observable macro variables by augmenting the state X with employment and CPI inflation. I also report the estimation results of three factor versions of the model. First, it is interesting to see whether the information the Fed cares about shows up in some other variable when the inertia factor is omitted, in which case $X = (\theta, s, v)$. Second, yields have started to look 'more Gaussian' in the 1990s (in terms of the tails of the histogram of yield changes). Constant volatility $\sigma_s = \sqrt{v}$ might thus be sufficient to capture yields.

The continuous specification of s implies that the short rate always jumps one-to-one with the target. Whether this assumption is sensible is hard to judge from, say, funds-rate data only (which is one of the reasons to work with an unobservable short rate). Some confidence in this assumption can be gained from the behavior of the funds rate around target changes in Figure 1, which is particularly evident for the year 1994. An easy extension is to let the spread jump by some fixed or state-independent random amount.

3.5 Solving for the Yield Curve

Arbitrage-free pricing now specifies an exogenous risk-adjustment in the form of a 'density process' ξ. Asset prices are then given by the conditional expected value of their payoff, weighted by ξ, and discounted at the riskless short rate r In particular, the time-t price of a zero-coupon bond price that matures at time T is

$$P(t,T) = E_t \left[\frac{\xi(T)}{\xi(t)} \exp\left(-\int_0^t r(u) du \right) \right], \tag{6}$$

where E_t denotes expectation given the information available to bond investors at time t. In a Lucas (1978) economy, for example, the term inside the expectation is just the marginal rate

of substitution of a representative agent. The weight ξ can be used to define a risk-neutral probability measure $\mathcal{Q}$ which satisfies $E_t\left[Z\xi(s)/\xi(t)\right] = E_t^{\mathcal{Q}}(Z)$ for any random variable Z known at time s for which this expectation exist. If such a risk-neutral measure $\mathcal{Q}$ exists, there is no arbitrage, at least under reasonable restrictions on trading strategies (Harrison and Kreps (1979), Harrison and Pliska (1981)). Conversely, the absence of arbitrage, and some technical conditions, implies the existence of a risk-neutral measure (Delbaen and Schachermayer (1994)).

The computation of a solution $P(t,T)$ to (6) can be performed only in a few special cases by hand. For example, if r is Gaussian under the risk-neutral measure $\mathcal{Q}$, then this just involves taking the expectation of a sum of Gaussians, which can be computed directly (Vasicek (1977)). For the general case in which X is a LQJD under $\mathcal{Q}$, the idea is first to guess that bond prices are given by the exponential LQ form

$$P(t,T) = \exp\left(c_0(t,T) + c_1(t,T)X(t) + X(t)^{\intercal}c_2(t,T)X(t)\right), \tag{7}$$

for some coefficients $c(t,T) = (c_0(t,T), c_1(t,T), c_2(t,T)) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^{N\times N}$ that depend on the particular ordering of deterministic jump dates between t and T. This guess is verified by computing $c(t,T)$ recursively using the method of undetermined coefficients, starting at the time T of maturity with the boundary condition $c(T,T) = 0$, implied by the fact that $P(T,T) = 1$, and from the assumption that D contains an open set. Proposition 1 in Appendix B.3 states this result.

The recursive computation of bond prices (and thus coefficients) iterates two steps. Roughly speaking, the first step (Lemma 1 in Appendix B.6) shows that if the bond price at the next deterministic jump date is an exponential LQ function in the state vector, as in (7), then the price of a bond "just before" the jump date is of the same form. The second step (Lemma 2 in Appendix B.7) demonstrates that if the bond price "just before" the next deterministic jump date is given by the exponential LQ form (7), then the price during the entire interim period between two deterministic jump dates solves a partial differential equation that also has an exponential LQ solution. Together, these two steps guarantee that for every t, the price $P(t,T)$ inherits the postulated form.

3.6 Risk Adjustment

The density process ξ solves $d\xi = -\xi\sigma_\xi dW$, where $W = (W_s, W_v, W_z)$, so that a standard Brownian motion $W^{\mathcal{Q}}$ under $\mathcal{Q}$ solves $dW^{\mathcal{Q}} = dW + \sigma_\xi dt$. The market prices σ_ξ of uncertainty for the Brownian motions are of the form

$$\sigma_\xi = \begin{pmatrix} \sigma_\xi^s(t) \\ \sigma_\xi^v(t) \\ \sigma_\xi^z(t) \end{pmatrix} = \begin{pmatrix} q_s\sqrt{v(t)} \\ q_v\sigma_v\sqrt{v(t)} \\ q_z \end{pmatrix} \tag{8}$$

leading to time-varying risk premia that are affine in the volatility factor v. The three factor versions of the model do not need some of the components in σ_ξ. In particular, if volatility

σ_s is constant, σ_ξ^s is constant as well. The parametrization of σ_ξ also captures risk aversion against target moves that are driven by s, v and z. This means that even without market prices of uncertainty for N_U and N_D, the intensities under the risk-neutral measure $\mathcal{Q}$ may differ from their values under the data-generating measure because of their state-dependence. With only 5 years of data, I chose to set the market prices of jump uncertainty for target-rate moves to zero. For example, λ_0^U is hard to estimate even without any risk adjustment.

4 Estimation Technique and Data

This section describes the simulation-based method used to approximate the joint likelihood function of the target, LIBOR and swap rates, which is not available in closed form. Moreover, it presents the approximation of the pricing map used in the estimation.

4.1 Estimation Problem

Let $f_X(\,\cdot\,, t|X_{\tilde{t}}, \tilde{t}; \gamma)$ denote the true density of the state vector X_t conditional on the last observation $X_{\tilde{t}}$ at some $\tilde{t} < t$. The parameter vector γ contains parameters describing the true distribution of X and parameters governing the market prices of uncertainty. This density involves the nonlinear stochastic intensities in (2). Let $p(\,\cdot\,, \gamma)$ denote the true mapping from factors to observed yields and the target for a given γ, in that $p(X_t, \gamma) = Y_t$, where Y_t is the vector of observables at time t: yields and the target rate θ_t. Assume that $p(\,\cdot\,, \gamma)$ can be inverted to obtain the factors as function $q(\,\cdot\,, \gamma)$ of the observables Y_t, in that $X_t = q(Y_t, \gamma)$.

Ideally, the model would be estimated by maximizing the likelihood of the observations over γ, which can be obtained by a change of variables from the conditional densities of the state variables. For example, the conditional density $f(\,\cdot\,, t \,|\, Y_{\tilde{t}}, \tilde{t}; \gamma)$ of Y_t given $Y_{\tilde{t}}$ at $\tilde{t} < t$ is given by

$$f\left(Y_t, t \,|\, Y_{\tilde{t}}, \tilde{t}; \gamma\right) = f_X\left(q(Y_t, \gamma), t \,|\, q(Y_{\tilde{t}}, \gamma), \tilde{t}; \gamma\right) \left|\nabla_Y q(Y_t, \gamma)\right|. \tag{9}$$

Three problems arise. First, the true density f_X of the state variables is not available in closed form. I therefore extend the simulated maximum likelihood (SML) method of Pedersen (1995) and Brandt and Santa-Clara (2001) to jump-diffusions (Section 4.2). Second, the true maps p and q are not available in closed form. In this high-dimensional setting, Monte-Carlo integration to recover $p(\,\cdot\,, \gamma)$ is prohibitively expensive. This roadblock is bypassed by using an approximating LQJD model, for which the Jacobian term in (9) can be calculated analytically. A hill-climbing procedure based on analytical derivatives inverts the map from states to LIBOR and swap yields numerically for each observation (Section 4.3). Third, an exact computation of yield coefficients for the approximating LQJD model is computationally intensive, so I employ a time-saving algorithm (Section 4.4).

4.2 Density Approximation (SML)

The conditional density of the likelihood function of the underlying state vector solves a partial differential-integral equation that has a closed-form solution only for a few special cases, such as Gaussian and square-root diffusions (see Lo (1988)). To overcome this problem, SML is used, which attains approximate efficiency similar to the efficient method of moments technique by Gallant and Tauchen (1996).[12] The density $f_X(\,\cdot\,, t \,|\, \tilde{x}, \tilde{t})$ of the state $X(t)$ conditional on the last observation $X(\tilde{t}) = \tilde{x}$ can be written, using Bayes' Rule and the Markov property of X, as

$$f_X(x, t \,|\, \tilde{x}, \tilde{t}) = \int_D f_X(x, t \,|\, w, t-h)\, f_X(w, t-h \,|\, \tilde{x}, \tilde{t})\, dw, \tag{10}$$

for any time interval h. (This is sometimes called the Chapman-Kolmogorov equation.) SML computes (10) by Monte-Carlo integration, replacing the density $f_X(\,\cdot\,, t \,|\, w, t-h)$ by the density of a discretization of X. The method is extended in Appendix C.1 to allow for jump-diffusions. Particular care needs to be taken to accommodate time-dependent stochastic intensities.[13]

4.3 Pricing-Formula Approximation

Modeling the target-rate with jump intensities defined by (2) introduces a form of non-linearity that takes the state vector outside of the LQJD class as mentioned in Section 3. The quality of an approximating pricing formula $Y_t = \tilde{p}(X_t, \gamma)$ that ignores the truncation by the max-operator in (2) depends crucially on how severely the positivity constraint on the intensities is binding (and on the average impact of hitting the constraint). The inverse $\tilde{q}(\,\cdot\,, \gamma)$ of the approximating map $\tilde{p}$ is defined in the obvious way. Let

$$D_+^{\gamma_0} := \{x \in D : \lambda_0^j + \lambda_X^j x \geq \gamma_0^j,\ j = U, D\}$$

denote the set of states at which the intensity formula is bounded below by a given constant γ_0^j, for $j = U, D$. The approximation $\tilde{q}(\,\cdot\,, \gamma)$ is likely to be better at states in D_+^0.

[12] EMM implements a simulated method of moments estimator with moments generated by the scores of an auxiliary semi-nonparametric (SNP) density. The SNP density is a Hermite expansion (with analytical scores) that approaches the true density as the degree of the polynomial increases. In the case of SML, the simulated moments are scores from the discretized model. An alternative approximately efficient estimator for is proposed by Singleton (2001), who computes explicit moments using the conditional characteristic function $\psi(\,\cdot\,)$ of X, defined by $\psi(u) = E_{t-1}[\exp(i\, u^\top X_t)]$. Efficiency is achieved by increasing the number of different values taken on by u with one moment associated with each choice of u. This estimator can also be used in the LQJD setting, as the characteristic function can also be obtained in closed form. While SML is used here as a potentially helpful alternative to EMM, the computational costs of explicit moments as in Singleton (2001) are prohibitive in the present seasonal setting.

[13] Like any simulation-based technique, SML is computationally intensive. Unfortunately, the application in this paper does not share the advantage of analytical gradients with Brandt and Santa-Clara (2001), because the map $p(X_t, \gamma)$ involves the numerical computation of ODEs that depend on the parameters. The numerical optimization procedure is therefore based on the Nelder-Mead simplex method, starting a gradient-based parameter search only after the simplex algorithm has collapsed.

Two estimation approaches can be taken. The first approach is simply to replace p in (9) by $\tilde{p}$ and obtain an estimator of γ by maximizing the total approximate likelihood

$$\prod_{(\tilde{t},t)\in I} \tilde{f}(Y_t, t \,|\, Y_{\tilde{t}}, \tilde{t}; \gamma) =$$
$$\prod_{(\tilde{t},t)\in I} f_X\left(\tilde{q}(Y_t,\gamma), t \,|\, \tilde{q}(Y_{\tilde{t}},\gamma), \tilde{t}; \gamma\right) \left|\bigtriangledown_Y \tilde{q}(Y_t,\gamma)\right|,$$

where I denotes pairs of successive observation times in the data set. Here, the true factor dynamics (captured by f_X) are combined with the inverse $\tilde{q}$ of the approximating map from factors to yields. Since there is no restriction here on the parameter space, the approach is labeled *unconstrained estimation*. The accuracy of this approximation of the likelihood can be assessed *ex post* by checking whether the functions p and $\tilde{p}$ are close at the estimated parameter $\hat{\gamma}$.

One might be concerned that, without *a-priori* restrictions on the parameter space, it is unlikely that a good approximation of p would obtain at the estimated parameter value. Even though this turned out not to be a problem in the application below, an alternative approach was also tried. Specifically, consider performing the *constrained estimation*

$$\max_{\{\gamma,\gamma_0\}} \prod_{(\tilde{t},t)\in I} \tilde{f}\left(Y_t, t \,|\, Y_{\tilde{t}}, \tilde{t}; \gamma\right) \tag{11}$$
$$\text{subject to} \qquad \tilde{q}(Y_t,\gamma) \in D_+^{\gamma_0}, \text{ for all } t \in I.$$

In words, the parameter space is restricted to contain only those parameters at which the observations are explained by a factor realization $\tilde{q}(Y_t,\gamma)$ for which the γ_0-constraint never binds. The special case that was tried in this paper is $\gamma_0 = 0$. Naturally these two problems typically deliver distinct estimators. Any such differences will be further discussed when the estimates are presented.

4.4 Coefficient Approximation

Time dependencies introduced by scheduled announcements, such as FOMC meetings and macro releases, immensely increase the computational burden associated with the solution of the approximating LQJD model for yields, and render almost impossible an estimation using data for long-maturity yields. For example, in order to evaluate the likelihood function, the 5-year swap rate needs to be computed for each observation in the sample. In the setup of Section 6, this takes 16 minutes on a SUN workstation.[14] In setups with only one type of scheduled announcement (at FOMC meetings), the coefficients are therefore computed using

[14]The computation simulates 9 coefficients (for each of the 8 state variables plus a constant) for 10 different bond maturities ($0.5, 1, \ldots, 4.5, 5$ years) for each of the 261 observations, using Runge-Kutta solutions of the ODEs for the coefficients.

the following approximation: the time until the next FOMC meetings is matched exactly only for the next-to-occur meeting. The subsequent meetings are assumed to be equally spaced over the year. For the maturities of the yields used in the estimation (6 months and above), the errors due to this approximation are virtually undetectable. In setups with more than one type of scheduled announcement (FOMC meetings and macro announcements), such an approximation is no longer accurate and is not pursued in this paper.

4.5 Data

The target rate is on average higher than short Treasury rates, because the Fed is operating in a market where loans are not collateralized, while tax effects and other factors further depress Treasuries. For example, the average daily target rate from 1994 to 1998 is 5.22%, while the 3-month T-bill and the 3-month LIBOR-rate averaged 5.06% and 5.44%, respectively. The empirical results reported here are therefore based on LIBOR and swap rates.[15] The sample period considered in the estimation is January 1, 1994 to December 31, 1998. The dates of FOMC meetings were obtained from the Board of Governors of the Federal Reserve. The FOMC meets eight times a year. Two of these meetings, the first and the fourth, extend over two days. In the past, if the Fed changed its target during one of these two-day meetings, the announcement was always made on the second of the two meeting days. The target-rate series used in this paper differs from the series in Datastream with respect to the timing of the target change during the two-day meeting of February 1994. Datastream assigns the change to the first meeting day (February 3), while the change was announced on the second meeting day (*The New York Times*, February 5, 1994, page 1, "Federal Reserve, Changing Course, Raises a Key Rate" by Keith Bradsher).

LIBOR data are from the British Bankers' Association, while swap rates are from Intercapital Brokers Limited. Both series are obtained through Datastream. LIBOR rates are recorded at 11 a.m. London time, while swap rates are recorded at the end of the UK business day. Target-rate changes are typically announced from 10 a.m. to 3 p.m. Eastern time.[16] This means that a move in the target on Tuesday, March 1, affects recorded LIBOR rates on Wednesday, March 2. The effect on recorded swap yields is not so precisely separable. The data sample is therefore constructed by using Thursday (London time) observations of LIBOR and swap yields, together with Wednesday (Eastern time) observations of the target rate. The asynchronous nature of the observations is ignored in the estimation. Whenever the respective day was a holiday, the observation of the previous business day was used.

The bond-pricing formula (7) extends as written to the case of LIBOR bonds, treating

[15] LIBOR-quality swap rates are minimally affected by credit risk because of their special contractual netting features, although they do trade at spreads to Treasuries that have, to this point, resisted a convincing explanation (see, for example, Collin-Dufresne and Solnik (2001)). For example, the 2, 5 and 10-year average swap rates were 6.08%, 6.48% and 6.79% during the post-1994 period, respectively, while the same maturities had average percentage yields in the Treasury market of 5.81, 6.13, and 6.34.

[16] For example: Sep 29, 1998 at 2:15 p.m., Oct 15, 1998 at 3:15 p.m., Nov 17, 1998 at 2:15 p.m.

r as a default-adjusted discount rate (Duffie and Singleton (1997)). The 6-month LIBOR rate is denoted $Y(t, t+0.5)$. An interest-rate swap is a contract between two parties to exchange fixed and floating coupons for a stipulated time, say τ years. One party receives a semi-annual floating payment in form of the 6-month LIBOR rate and pays in exchange a fixed coupon rate, the swap rate. At the initiation of the swap contract, the swap rate is set so that the value of the swap contract is zero. For simplicity, swap rates $Y(t, t+\tau)$ are treated as par-coupon rates on LIBOR-quality bonds of the same maturity, putting aside the distinct institutional features and differences in default risk of LIBOR and LIBOR-swap markets, so that

$$Y(t, t+\tau) = \frac{2\,(1 - P(t, t+\tau))}{\sum_{j=1}^{2\tau} P(t, t+0.5\,j)}. \tag{12}$$

5 Estimation Results

The same set of yields (6-month LIBOR, 2 and 5-year swap) and the target are used for all estimations, creating the need to break the stochastic singularity arising from the exact map of three factors in four observed variables in the lower-dimensional systems. I therefore assume that the 2-year swap rate is observed with measurement error.[17] As this section investigates the properties of the estimated models, the concept of *model-implied factors* will be needed. These are obtained by inverting the map from factors to the target rate, and to those yields that are assumed to be observed without error at the SML estimates. The map from factors to yields is given by the pricing formulas (12) from the approximating model.

5.1 Accuracy of the Approximating Model

As the true state process X is not actually of the LQJD class, because of (2), it is important to study the accuracy of the approximating model. From the swap-yield formula (12), one can see that it is sufficient to investigate the approximation accuracy for zero-coupon yields. Zero-coupon yields $Y_0(t, T)$ for each observation t in the sample implied by the true (nonlinear) model can be computed with Monte Carlo integration. Consider simulating S paths of the short rate for times $i = t+h, t+2h, \ldots, T-h$ starting with the model-implied state x at time t. The time t yield of a zero-coupon bond maturing at time T is then

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

[17] The measurement error of the 2-year swap rate turns out to be persistent, with a weekly autocorrelation coefficient that varies across models from 0.95 to 0.98, and a standard deviation that varies from 12 to 21 basis points. This is comparable to other yield-curve estimations (see, for example, Duffie and Singleton (1997)).

where

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

These calculations are performed using $S = 10,000$ and $h = 1/365$. FOMC meeting days are additionally subdivided into 30 intervals. Given these choices, the standard errors of the Monte-Carlo approximation of the true yields for even the 5-year yield are sufficiently small, from 0.93 to 1.79 basis points.

These zero-coupon yields implied by the true model can be compared to the yields from the approximating model that are based on the same model-implied state vector. Table 1 shows that the mean absolute approximation error is around 1-3 basis points, with a standard deviation of 1-2 basis points for the constrained model without inertia factor z and an even smaller error for the full model. The approximation error for these specifications is thus of similar magnitude as the bid-ask spread of swaps. The unconstrained estimates of the model without z and all other 3-factor versions produce about 5 times the approximation error, which seems too large to be acceptable.

5.2 Cross-Sectional Fit of the Approximating Model

An important measure of cross-sectional out-of-sample fit for any yield-curve model is the pricing error made when predicting yields that were not included in the estimation. Pricing errors are defined as the difference between actual yields and the model-implied yields computed by inserting the model-implied factors into the swap formulas (12). Model-implied yields are based on the approximating model, so that the pricing errors not only reflect model misspecification, but also approximation error. Table 2 reports the average absolute pricing errors and their sample standard deviations for the base-case models and for the Dai and Singleton (2000, DS) model as a point of reference. The DS parameters are estimated with weekly data on the same LIBOR and swap rates with the exception of using the 10-year instead of the 5-year swap over a sample period that only partially overlaps with the sample used in this paper (April 3, 1987 to August 23, 1996). The parameters have not been reestimated.

The pricing errors of the unconstrained full model and its version without stochastic volatility are lowest among these. The model without stochastic vol matches the short end of the yield curve extremely well with average absolute errors of around 6 to 13 basis points. In addition to matching the short end well, with 11-26 bp errors, the full model produces low errors, of around 2 bp, at the long end of the curve. As can be seen from the table, the latter model does better than the DS model, especially at the short end. The incorporation of the target as a fourth factor appears to provide a manageable way of fixing the short end of the yield curve.

Table 1: Approximation Errors due to 'Negative Intensities' (in Basis Points)

Maturity		W/o Stoch. Vol		W/o Inertia		Full Model	
		Con.	Unc.	Con.	Unc.	Con.	Unc.
6-mth	mean abs.AE	4.08	8.56	2.85	11.32	2.60	2.11
	std of abs.AE	2.73	9.26	2.20	10.69	1.91	1.11
	average SE	0.53	0.38	0.46	0.29	0.59	0.67
2-yr	mean abs.AE	11.43	20.99	2.62	19.43	2.17	1.73
	std of abs.AE	8.78	18.59	1.50	18.18	1.36	0.85
	average SE	0.93	0.66	0.85	0.59	1.07	1.13
5-yr	mean abs.AE	37.74	28.92	1.76	9.37	1.54	1.86
	std of abs.AE	20.16	20.74	0.72	9.30	0.81	0.73
	average SE	1.19	0.93	1.28	1.07	1.64	1.79

NOTE: This table presents summary statistics about the approximation errors in basis points made by the approximating models over the sample January 1, 1994 to December 31, 1998. "W/o Stoch. Vol" refers to (5) with constant σ_s, while "W/o Inertia" takes out the z process so that $X = (\theta, s, v)$. "Con." refers to the constrained estimation as in (11), "Unc." to the unconstrained counterpart. Due to the seasonality introduced by FOMC meetings, the approximation errors in this setup depend on time t even for a given value of the state vector. The table therefore reports the mean average absolute approximation error $|Y(t,T) - Y_0(t,T)|$ and its standard deviation over the sample (first and second row). The table also reports the average standard errors of the Monte Carlo approximation of true yields $Y_0(t,T)$ (third row). These 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 i.i.d. population of random variables $\exp(\sum_{i=t}^{T-h} \hat{r}_i[s]h)$. The table reports the average standard errors over the sample.

Taking out the z variable produces overall higher pricing errors (ranging from 8 to 18 bp) when the model is evaluated at the constrained parameter vector, and about 1-7 bp higher than that for the unconstrained parameter.[18] Given the small pricing errors of unconstrained full model, this will be the model I will focus mostly on in the following analysis.

5.3 Interpretation of Model-Implied Factors

The spread s represents *money-market noise* which only affects short maturities. Table 3 shows that this feature is shared by all models (as defined in Section 5). Under the data-generating probability measure, the short rate quickly reverts to the target, which in turn reverts to a parameter $\bar{\theta}$ that is fixed to its sample mean, 5.22%. The speed κ_s of

[18] While pricing errors are a measure of first moments that depends on the realized path of yields through the use of model-implied factors, a history-independent check can be obtained by simulating 20,000 samples of weekly yields. The average 6-month LIBOR, 2 and 5-year swap rates in these simulated samples are 5.59, 5.76 and 5.89%, respectively, which shows better how the high persistence of yields makes it difficult to match first moments.

Table 2: Pricing Errors for Yields not used for Fitting (in Basis Points)

		W/o Stoch. Vol		W/o Inertia		Full Model		DS Model
		Con.	Unc.	Con.	Unc.	Con.	Unc.	
1 mth	mean	13.5	12.53	18.48	25.41	33.57	25.72	237.34
	std	11.07	11.29	11.59	13.96	16.69	16.36	115.48
3 mth	mean	7.90	7.48	11.85	15.48	16.74	11.00	66.17
	std	6.22	6.33	6.65	7.71	7.41	6.75	30.72
12 mth	mean	6.67	6.84	9.90	12.44	4.72	3.22	9.93
	std	6.48	6.58	7.17	8.77	3.53	2.92	4.83
3 year	mean	10.68	9.47	14.96	16.94	8.46	1.76	6.41
	std	5.89	5.66	8.79	9.03	2.35	1.13	1.54
4 year	mean	6.40	5.76	7.98	8.81	10.14	1.78	5.70
	std	3.17	3.08	4.55	4.51	2.45	1.02	1.51

NOTE: This table presents the mean and the standard deviation of the absolute value of the pricing error in basis points over the weekly data sample from January 1, 1994 to December 31, 1998 made by the approximating model and, as a reference, for the DS model (at their parameter estimates). The definition of models is as in Table 1. Using their sample period, DS report *mean pricing errors* for weekly 3, 5 and 7 year swaps rates of -11.3, 16.9 and -12.7 basis points with standard deviations of 9.6, 16.5 and 10.1 basis points.

mean-reversion is highest in the full model, implying a weekly autoregressive coefficient of $\exp(-\kappa_s/52) = 0.83$ and a half life of shocks to the spread of less than 1 month. Under the risk-neutral measure, spread shocks die out fast as well, which implies that they only matter for short yields. Figure 3 computes, for the full model, the linear dependence of zero-coupon yields on all factors, as a function of maturity. These yield coefficients can be interpreted as responses of yields to the various shocks because the shocks are uncorrelated. The response of yields to changes in the spread s is monotonically decreasing with maturity, which means that s is a "slope factor" in the language of Litterman and Scheinkman (1993).

Shocks to the target persist under both probability measures which leads to positive autocorrelation in the target level. This will be referred to as *interest-rate smoothing.* Target shocks therefore mostly affect short yields and less long yields, as can be seen from Figure 3. This monotone reaction of yields is usually called liquidity effect. In other words, both s and θ are "slope factors", but act on different parts of the yield curve, as the impact of target increases dies off more slowly with maturity than do the impacts of shocks to s.

The Fed conducts monetary policy by slowly pulling the target (and thus the short rate) towards a desired rate that is mostly captured by z (under both probability measures). This can be seen from Table 4 which documents that arrival rates of Fed moves are most highly correlated with the inertia factor z. Positive shocks to z increase the likelihood of a target move not only at the next FOMC meeting, but also at subsequent meetings, because their half live is 1 year (again, under both measures) from Table 3. This means that certain

Table 3: Estimated Parameters of Yield-Curve Models with Target

	W/o Stoch. Vol		W/o Inertia		Full Model	
	Con.	Unc.	Con.	Unc.	Con.	Unc.
κ_s	1.28	1.56	2.86	3.33	6.44	9.75
	(2.53)	(4.10)	(3.90)	(4.20)	(3.83)	(4.69)
κ_v	–	–	0.17	0.1296	0.12	0.04
	–	–	(3.74)	(3.80)	(0.41)	(0.42)
κ_z	0.28	0.29	–	–	0.47	0.72
	(2.69)	(3.03)	–	–	(4.73)	(4.34)
θ	0.0522	0.0522	0.0522	0.0522	0.0522	0.0522
	–	–	–	–	–	–
$\bar{v}$	–	–	0.000045	0.000059	0.000228	0.000415
	–	–	(5.69)	(4.53)	(0.51)	(1.07)
σ_s	0.0000712	0.0000799	–	–	–	–
	(51.39)	(12.50)	–	–	–	–
σ_v	–	–	0.0058	0.0065	0.0542	0.0899
	–	–	(2.04)	(1.37)	(0.72)	(0.78)
λ_0^U	758.2	338.6	341.7	−323.3	1160.3	273.7
	(0.11)	(0.40)	(0.03)	(−0.02)	(0.15)	(0.55)
λ_s	6769	7582	20132	18685	6117	7267
	(1.48)	(2.25)	(2.31)	(2.14)	(2.32)	(1.86)
λ_θ	−5269.7	−4876.7	−12638.3	−14795.8	−8738.6	−9408.9
	(−3.90)	(−4.30)	(−7.48)	(−9.23)	(−6.53)	(−4.18)
λ_v	–	–	17831000	18611780	−89192	548315
	–	–	(2.55)	(2.26)	(−0.60)	(1.63)
λ_z	123	119.5	–	–	234.8	237.6
	(18.74)	(18.63)	–	–	(21.79)	(17.87)
q_s	70.12	70.29	−257.53	−241.32	−89.93	−47.62
	(3.27)	(8.67)	(−5.29)	(−4.79)	(−3.22)	(−2.90)
q_v	–	–	499.9	1051.16	−10186	−2537.5
	–	–	(0.05)	(0.12)	(−1.19)	(−0.81)
q_z	−0.2091	−0.2132	–	–	0.2553	0.1126
	(−132.73)	(−33.36)	–	–	(27.82)	(0.18)
σ_M	0.00137	0.00213	0.00116	0.00114	–	–
	(15.12)	(14.77)	(17.54)	(18.14)	–	–
ρ_M	0.961	0.955	0.982	0.984	–	–
	(58.64)	(56.13)	(78.93)	(88.11)	–	–

NOTE: This table reports the SML parameter estimates and t-ratios (in brackets) obtained with $S = 2500$, $h = \frac{1}{M}\frac{1}{365}$, $M = 1$, $M_s = 30$ and weekly observations of the target rate, 6-month LIBOR, 2 and 5-year swap rate from January 1, 1994 to December 31, 1998. "Full model" refers to equations (1), (2), (5) and (8). λ_θ, for example, denotes the slope coefficient of θ in λ_X. "W/o Stoch. Vol" refers to (5) with constant σ_s, while "W/o Inertia" takes out the z process so that $X = (\theta, s, v)$. "Con." refers to the constrained estimation as in (11), "Unc." to the unconstrained counterpart. σ_M is the standard deviation of the measurement error contaminating observations of the 2-year swap rate and ρ_M is its autocorrelation.

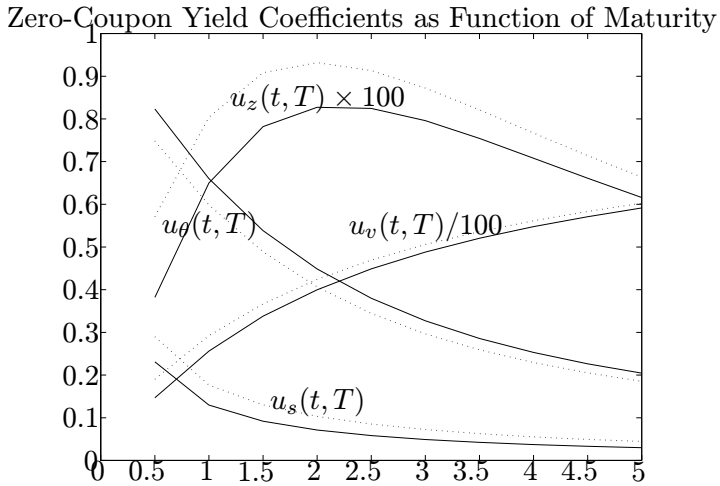


Figure 3: Zero-Coupon Yield Coefficients as a function of time to maturity ($T - t = 1, 2, \ldots$ years) taken from the (unconstrained) full model: $Y_0(t) = u_0(t,T) + u_s(t,T)s(t) + u_\theta(t,T)\theta(t) + u_v(t,T)v(t) + u_z(t,T)z(t)$ at two typical dates t. The solid line is at an FOMC meeting. The dotted line is at the day after the meeting.

shocks to the economy affect the target and thus short-term rates only with some delay, and thereby induce positive autocorrelation in target changes. This positive autocorrelation in changes is labelled *policy inertia*. Figure 3 shows that the anticipated cumulative effect of subsequent target moves on yields is highest for maturities around 2 years. At sufficiently long maturities, beyond 2 years, mean reversion in the target and the inertia factor causes shocks to have smaller and smaller impacts on longer and longer rates. The net effect is a hump-shaped coefficient on z, which means that z is a "curvature factor" in the language of Litterman and Scheinkman (1993).

Shocks to volatility v affect yields at all maturities, reflecting the strong persistence of volatility (under both measures). In this sense, v is a "level factor".

Table 4: Which Factors drive the Probability of Target Moves ?

	W/o Stoch. Vol		W/o Inertia		Full Model	
	Con.	Unc.	Con.	Unc.	Con.	Unc.
s	.44	.53	.60	.55	.40	.48
θ	−.33	−.33	−.48	−.48	−.33	−.36
z	.85	.83	–	–	.86	.71
v	–	–	.65	.67	−.09	.26

NOTE: To obtain a measure of importance of a factor for the stochastic intensities of policy events, this table shows the correlation between changes in the model-implied factors (s, θ, v, z) and changes in the function $\lambda_0 + \lambda_s s(t) + \lambda_\theta \theta(t) + \lambda_v v(t) + \lambda_z z(t)$ for the weekly sample from January 1, 1994 to December 31, 1998. The definition of models is the same as in Table 1.

Table 5: Correlations of Model-Implied in Factors, Yields and Target

Model		W/o Stoch. Vol		W/o Inertia		Full Model			LIBOR & Swaps			Target
		r	z	r	v	r	z	v	6-mth	2-yr	5-yr	θ
W/o Stoch. Vol	r	1							.56	.26	.14	.66
	z	.01	1						.44	.89	.97	.07
W/o Inertia	r	.85	−.18	1					.57	.13	−.01	.75
	v	.07	.99	−.09	1				.55	.93	.99	.12
Full Model	r	.67	−.24	.77	−.16	1			.54	−.03	−.07	−.12
	z	24	.63	.12	.64	−.18	1		.44	.81	.65	.13
	v	-.08	.78	−.19	.78	−.01	.01	1	.37	.55	.76	−.03
DS	r	.07	−.57	.21	−.54	.57	-.90	.05	−.05	−.62	−.50	−.26
	θ	.23	.82	.10	.85	−.16	.93	.34	.63	.96	.86	.29
	v	−.15	−.10	−.26	.62	.02	−.19	.97	.19	.33	.59	−.21

NOTE: This table computes the correlation of the first differences of model-implied factors (r, z and v) from the unconstrained estimations, the model-implied factors (r, θ, v) of the "$A_1(3)_{DS}$ model" by Dai and Singleton (2000) (at their estimated parameter vector), the 6-month LIBOR rate, the 2 and 5-year swap rates, and the target rate θ over the sample January 1, 1994 to December 31, 1998. The definition of models is the same as in Table 1. All correlations with the target rate are computed using the subsample of FOMC meetings.

The comparison of model-implied variables across models in Table 5 shows that the latent variables z and v roughly correspond to the stochastic mean "θ" and volatility "v" factors of the DS model, as can be seen from their correlation of 0.93 and 0.97, respectively. The commovements with yields indicate that z behaves much like the 2-year swap rate, and that v is related to the 5-year swap rate. The two models imply, however, very different short rates. The sample mean of the model-implied r in the DS model is -0.46%, while the average r in the full model is 5.02%. The full model implies that r is closely related to the short end of the yield curve, while the DS model produces a short rate that behaves like the slope at the short end of the yield curve. This can be seen from the correlations between the model-implied short rates r from the two models and the LIBOR rate (0.54 for the full model and -0.05 for the DS model), and the difference between the 2-year swap and the LIBOR rate (0.58 for the full and 0.86 for the DS model).

Similar to the full model, the short rate r in the three factor versions is correlated most highly with the shortest yield in the estimation. In the model without inertia factor z, r is less related to longer yields than in the version without stochastic volatility v. This can also be seen from the estimated mean-reversion parameter κ_s, higher for the model without z than for the one without v, and its correlation with target-rate changes on FOMC meetings, which is also higher in the model without z. This may be explained by the fact that, for a nonzero market price q_s of uncertainty, the model without z has the additional flexibility of allowing r to revert under $\mathcal{Q}$ to the continuous variable v, while the conditional mean of r between

FOMC meetings under the risk-neutral measure is constant in the model with constant vol. The second latent factor in these models behaves much like the longest yield from Table 5. They mean-revert at about the same slow rate and are in fact highly correlated, but almost uncorrelated with their short rate r. This may indicate that the role of v as conditional second moment of r is dominated by its importance of setting the arrival intensity of policy moves.

5.4 Snake-Shaped Volatility Curve

Figure 4 shows the standard deviation of yield changes as a function of maturity, the so-called term structure of volatility (or 'vol curve'). The vol curve is "snake-shaped," in that volatility is high at the very short end, declines until maturities about 3-6 months, after which it has a "hump" at a maturity of 2 years. The graph also shows the vol curve in simulated data from the full model, which reproduces the overall "snake-shaped" pattern quite well. The "head of the snake" at the short end is generated by money-market shocks, shocks to the spread, which only affect short yields. The "back of the snake", the hump at 2 years, has already been documented by Litterman, Scheinkman and Weiss (1988).

Informal accounts (for example, Fleming and Remolona (1999)) conjectured that monetary policy is responsible for the hump. This claim is here validated in the sense that the source of the hump are shocks to the inertia factor z which only enter the model through their impact on the arrival rates of policy moves. Rates at medium maturities such as 2 years respond immediately to the anticipated cumulative effects of z-shocks on the target, and thus have greater volatility then do the Fed-dampened short rates. Mean reversion in the target then leads to smaller volatilities for longer maturity yields. This means that the hump in the yield-coefficient on the inertia factor translates into a hump in the vol curve. The volatility in the data is higher around FOMC meetings, especially for short maturities. This seasonality is also present in the model, where it is generated by shocks to the target rate, which happen mostly at FOMC meetings. The model overstates this seasonality around FOMC meetings, because shocks to the target do not die out fast enough in the estimated model.

5.5 Implied Discrete-Choice Model of Policy

From each of the yield-curve setups, it is possible to derive a discrete-choice model in which, at each FOMC meeting, the Fed is viewed as randomizing over three possible policy choices: up move, down move or no move. The conditional probability of a particular choice at the FOMC meeting at t depends on the state "right before" t and is obtained from its empirical frequency in a simulated sample of size $S = 10,000$ that is generated by simulating forward in steps of one day's length starting at the actual value of the implied state at the last observation. These conditional probabilities from the full model are plotted in Figure 5 for each FOMC meeting. (Outside of FOMC meetings, the discrete-choice model assigns a small and constant probability to Peso events.) The conditional likelihood of moves up is

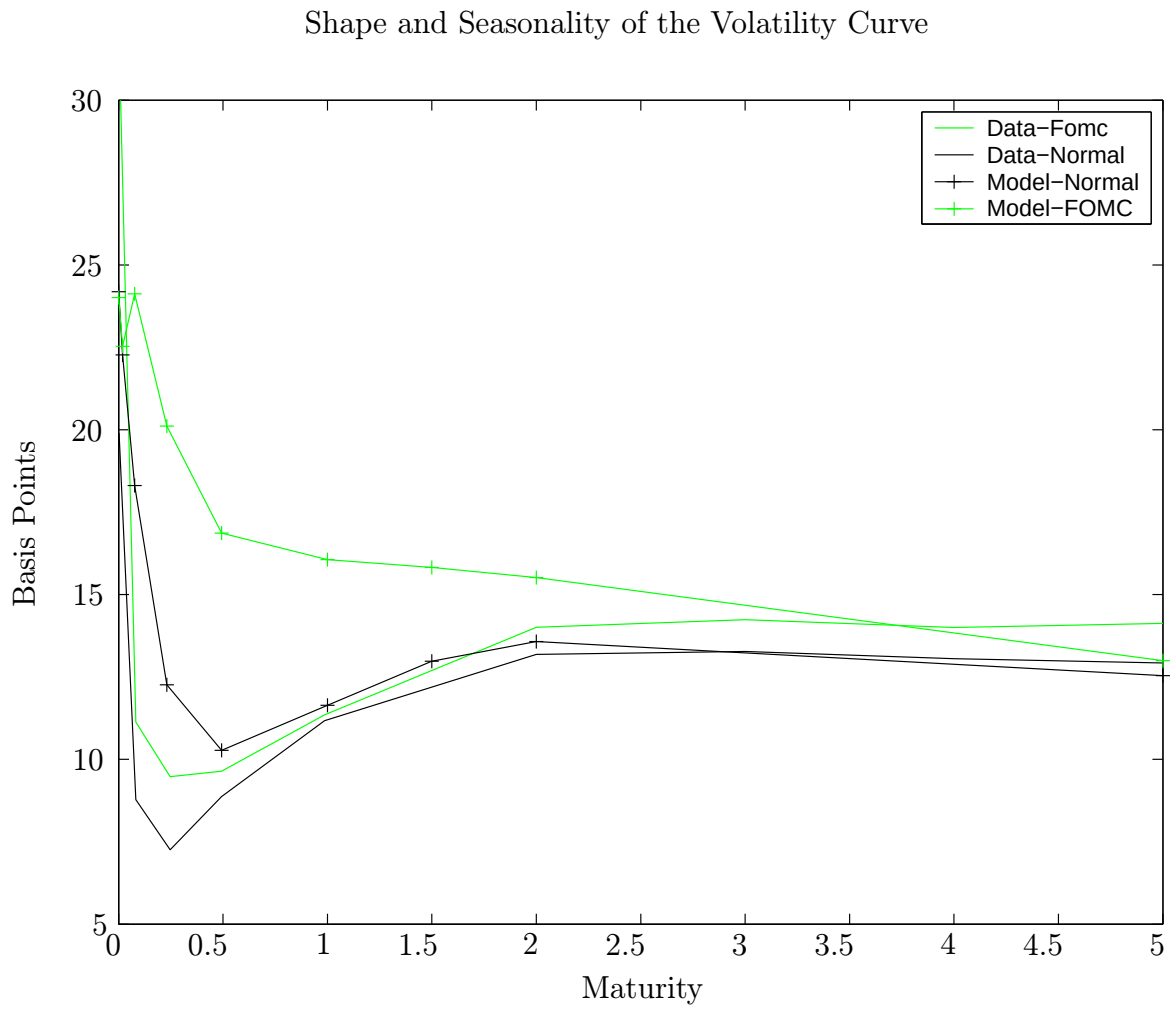
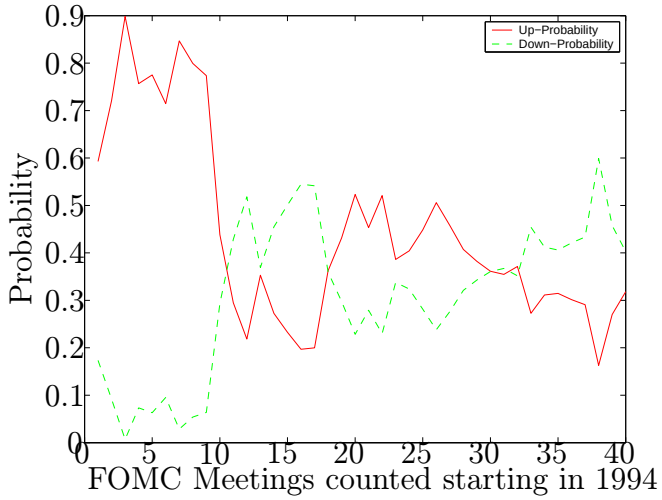


Figure 4: The graph shows the volatility curve during weeks of FOMC meetings (gray line) and the remaining weeks (black line) in the data (without +) and in a simulation of the full model (with +). The volatility curve is measured as standard deviations of yield changes. The weekly data are Wednesday observations from 1994 to 1998 on the target rate, overnight repo rate, and Thursday observations on the 1, 3, 6, 12-month LIBOR and 2, 3, 4, 5-year Swap Rates. Standard-error bounds around the volatility estimates in the actual yield data are computed in the following table with GMM using 5 Newey-West lags. The simulations for the simulated curve start with $S = 20,000$ initial states $\hat{X}_0$ that are obtained by simulating the state dynamics for 10 years, starting at the unconditional mean. Each day is subdivided into 2 intervals and each FOMC meeting is subdivided into 30 intervals. Given $\hat{X}_0$, S different samples of yields are simulated based on the actual FOMC calendar.

Standard Errors around Volatility Estimates (in Basis Points)

	Repo	LIBOR Rates				Swap Rates			
	Overnight	1 mth	3 mth	6 mth	12 mth	2 yr	3 yr	4 yr	5 yr
Data-	19.90	8.78	7.26	8.87	11.17	13.18	13.27	13.05	12.92
'Normal'	(2.01)	(1.53)	(0.86)	(0.84)	(0.79)	(0.75)	(0.76)	(0.70)	(0.67)
Data-	31.72	11.13	9.47	9.64	11.35	14.01	14.24	14.00	14.12
Fomc	(2.21)	(1.69)	(1.37)	(1.88)	(2.12)	(1.90)	(1.77)	(1.75)	(1.83)

Figure 5: Conditional Probabilities of Target Moves at each FOMC Meeting since 1994



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.

To see whether these conditional probabilities provide a good description of target dynamics, Table 6 compares forecasts by the model-implied discrete choice to forecasts based on alternative ways to describe target-rate moves. A forecast is taken to be the alternative with the highest conditional probability. As there have been only 7 increases and 5 decreases in the target at FOMC meetings over the sample period 1994-1998, these results suffer from small-sample noise, and are not intended to offer a serious forecasting comparison. They do provide, however, a device that helps to understand the implications of the model regarding Fed moves. The standard reference setup, usually labeled 'constant probability model,' is a version of the discrete-choice model in which the conditional probabilities are set equal to their empirical frequencies. For target moves, these frequencies are small (7/40 for "up" and 5/40 for "down"), so that this version always forecasts that the Fed is not going to change the target. In other words, this specification generates the same forecasts as a random walk for the target, and are reported under 'No Change.' A second reference model that seems useful is a random walk for the first differenced target; its forecasts are reported under 'Same Change.' For example, the first column of Table 6 indicates that of the 7 target-rate increases that occurred at FOMC meetings, the 'Same-Change' model would have predicted 2 correctly, while it would have forecasted no move in the remaining 5 cases.

The forecasts made by the unconstrained full model vastly outperform those of the constrained one in terms of the overall percentage of correctly forecasted Fed moves (75% and 30%, respectively). The reason is that the constrained model is characterized by large prob-

Table 6: Forecasting Evaluation of Model-Implied Target Model

	Same Change				No Change			
Predicted Actual	up	no	down	total	up	no	down	total
up	2	5	0	7	0	0	0	0
no	5	20	3	28	7	28	5	40
down	0	3	2	5	0	0	0	0
correct	2	20	2	24	0	28	0	28
total	7	28	5	40	7	28	5	40
% correct	28.57	71.42	40	60	0	100	0	70
	Unc. Full Model				Con. Full Model			
Predicted Actual	up	no	down	total	up	no	down	total
up	4	2	0	6	7	23	0	30
no	3	26	5	34	0	0	0	0
down	0	0	0	0	0	5	5	10
correct	4	26	0	30	7	0	5	12
total	7	28	5	40	7	28	5	40
% correct	57.14	92.86	0	75	100	0	100	30

NOTE: The sample used in this table is January 1, 1994 to December 31, 1998. During this time, there have been 40 FOMC meetings, 8 moves up in the target (1 outside of an FOMC meeting) and 6 moves down (1 outside of an FOMC meeting). This means that in a constant probability model, the estimated probability of a move up and down is 7/40 and 5/40, respectively. Forecasting a particular choice (up, down, no) is defined as the alternative with the highest probability. As to the 2 changes outside of FOMC meetings, all models would have missed them, so they are not part of the table.

abilities of nonzero moves. (This is also true for the three factor versions of the model). For example, the constrained model is not able predict a zero target-rate move correctly. It is worth noting, however, that the constrained model never misses the direction of the move, leading to a perfect score in forecasting nonzero moves. In other words, conditional on a Fed move, this model always predict the right sign of a move. The unconstrained model produces forecasts that are also more accurate than those of the reference models in terms of the overall correct forecasting percentage: the full model predicts 75% of the target moves on FOMC meetings correctly, compared to only 60% using the 'Same-Change' model and 70% using the 'No-Change' model.

5.6 Model-implied Policy Rule

Policy rules are structural equations that specify the map from a set of variables to the policy instrument of the central bank. Recursively identified VARs, for example, typically contain one equation describing the data-generating process of the fed funds rate, that can be interpreted as a policy rule plus some orthogonal monetary policy shock (see Christiano, Eichenbaum and Evans (1998)). The identifying assumption that will be used here is that

the Fed reacts to the value of the state 'just before' the FOMC meeting. The resulting *high-frequency rule* of the Fed can be computed as the conditional expected value of the target as a function of model-implied factors based on the estimated SML parameters.

For a given parameter value, the rule can be backed out by the staff using daily and even higher frequency data. The rule thereby avoids the criticism by Orphanides (1998) who argues that policy-rule plots regularly shown by the Fed staff to the FOMC at the beginning of each meeting cannot be based on right-hand-side variables that are yet to be released. With weekly observations, the model-implied policy rule at time t is

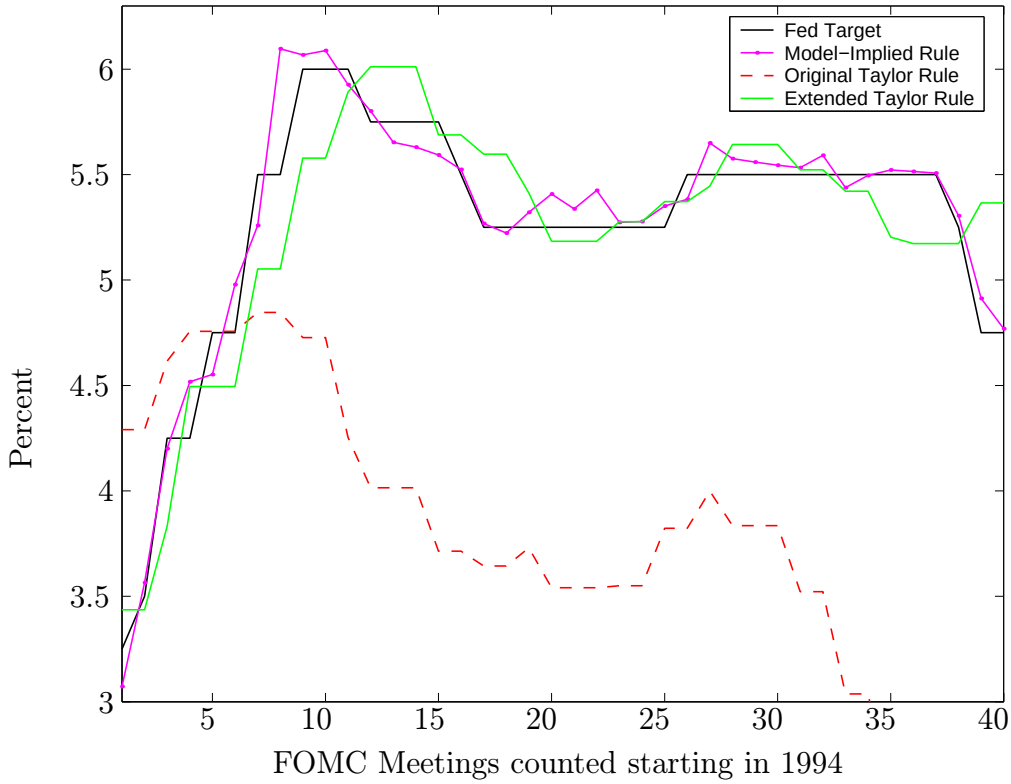
$$E_u\,[\,\theta(t)\,] = 0.36 + 0.10\,s(u) + 0.87\,\theta(u) + 7.51\,v(u) + 0.0033\,z(u), \tag{13}$$

where $u = t - 1/52$. According to this policy rule, the Fed reacts mainly to information contained in the yield curve. This statement is still true after employment and CPI inflation are added to the model (see Section 6). The most important variable in the rule is the inertia variable z which was shown to be closely related to the 2-year yield (Table 5). Another important variable in (13) is the past target, which is usually referred to as interest-rate smoothing. While variables such as volatility v and the inertia variable z may seem unfamiliar as arguments of a policy rule, they act as sufficient statistics for macroeconomic variables the Fed cares about. The small spread coefficient shows that the Fed does not care much about money-market noise which causes short-lived deviations of the short rate from the target.

Figure 6 compares the model-implied rule to the original Taylor rule and an estimated *extended Taylor rule*. Both Taylor-type rules use the quarterly averaged fed funds rate as left-hand side variable. The original Taylor rule uses the fixed coefficients proposed by Taylor (1993) on the two right-hand side variables: π four-quarter average inflation, computed using the GDP deflator, and the output gap y, computed as the percentage deviation of real GDP from its trend (based on a Hodrick-Prescott filter). The extension adds the lagged fed funds rate to the right-hand side.[19] To mimic the decision process of the Fed, the graph plots the policy rule for each FOMC meeting given its value in the quarter the meeting took place, leaving us with 40 data points. 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 actual target and the value of the target prescribed by the policy rule. For the original Taylor rule, the estimated Taylor (no lagged fed funds rate), the extended Taylor and the model-implied rule, the mean absolute difference is 162, 41, 22 and 10 basis points, respectively. The reason for the better fit of the model-implied rule is that the model-implied state variables, especially the inertia factor, commove with changes in the state of the economy to which the Fed reacts. This makes the inertia factor seem like it is 'anticipating' target moves, while the Taylor-based rules only catch up slowly with the target.

[19] The intercept of the original Taylor rule is 3, the coefficients on π and y are 1.5 and 0.5, respectively. The estimated rules are computed using quarterly data from 1994:1 to 1998:12. In the extended Taylor rule, the estimated constant is -0.3602, the coefficients on π, y and the lagged fed funds rate are 0.4968, 0.3411 and 0.9105 with standard errors 0.7473, 0.1907, 0.2251 and 0.0861, respectively, calculated with 2 Newey-West lags. The R^2 is 91%, which is reduced to 23% in the case of an estimated Taylor rule (without lagged fed funds rate). The data was obtained from the Federal Reserve Database.

Figure 6: Model-Implied Policy Rule together with Taylor Rules since 1994



6 Linking Policy Inertia to Macro Shocks

As the Fed reacts to macroeconomic variables (taken to be nonfarm payroll employment and CPI inflation) when fixing its target, expectations about future macro variables matter for current long yields. Whenever these expectations are updated, such as at macroeconomic releases, one would expect long yields to react. This conjecture can be verified in an unrestricted regression of yield changes on 'macro surprises', defined as the difference between actual and analyst forecast. For the case of employment releases, Figure 8 shows that the slope coefficient of this regression as a function of yield maturity has a hump at 2 years and is significant. In the following, I explore whether the information contained in the policy inertia factor of the yields-only model in the last section can be linked to macro shocks in a setting that respects the exact timing of analyst forecasts and macroeconomic releases.

6.1 Analyst Forecasts and Actual Data

Employment and CPI releases are made by the Bureau of Labor Statistics. Employment releases are at 8:30 a.m. on the first Friday of each month,[20] while CPI figures are released about two weeks after the end of the reference month, also at 8:30 a.m. This means that the LIBOR recorded at 11 a.m. London time is affected by a macro release on the preceding day, while swap rates recorded at the end of the London business day react on the same day as the macro release. The actual and released CPI and nonfarm payroll employment (NPE) series are from Money Market Services (MMS). The raw series obtained from MMS are the monthly percentage change in the CPI and changes in nonfarm payroll employment in thousands. The CPI series is multiplied by 1200 to obtain the annualized inflation rate, and changes in employment are divided by 100 (to obtain a series that is similar in magnitude to CPI inflation). These series of actuals are the numbers that were released on that particular date. MMS collects data on analyst forecasts each Friday prior to the actual release from about 40 money-market managers and reports their median forecast. These analyst-forecast data have been used in most studies of release surprises (for example, Balduzzi, Elton, and Green (1998), Li and Engle (1998), Fleming and Remolona (1999)). The monthly time series of changes in nonfarm payroll employment (NPE) and CPI inflation (CPI) for the sample period 1994-1998 contains only 60 monthly observations. Evidence about the macro dynamics will therefore be collected using all available data from MMS, which started surveying NPE forecasts[21] in January 1985.

6.2 Dynamics of Employment and Inflation

Over the sample period 1985:6 to 1998:12, it is not possible to outperform analyst forecasts in the mean-squared-error sense with one-step-ahead forecasts of univariate or bivariate ARMA specifications (even conditioning on past target values). The errors of analyst forecasts are positively correlated with those of time-series models, but this correlation is far from perfect. For instance, the sample correlation coefficient is at most 0.65 for the CPI and 0.85 for NPE. A reason for the relatively low correlation between analyst errors and model errors is the oversimplified informational structure assumed by the low-dimensional time-series model. When forecasting, actual investors are able to condition on a wealth of state variables. I thus explore a state-space model of macro variables and analyst forecasts that introduces latent variables summarizing this conditioning information.

The first step in setting up this system is to check for the unbiasedness of analyst forecast $m_F = (m_F^{CPI}, m_F^{NPE})$ of the vector $m = (m^{CPI}, m^{NPE})$ of CPI and NPE. I test for each

[20] The relevance of this variable can, for example, be seen from the Minutes of past FOMC meetings. In six out of eight FOMC meetings in 1996, the Board's discussion of the "economic and financial outlook" started with a general overview of the state of the economy and then immediately turned to the value of nonfarm payroll employment.

[21] The extension of the sample period to pre-1994 is justified if the precise timing of policy events (target moves at FOMC meetings versus moves at random business days) does not matter for how policy impacts monthly macro series, an assumption which seems to be reasonable.

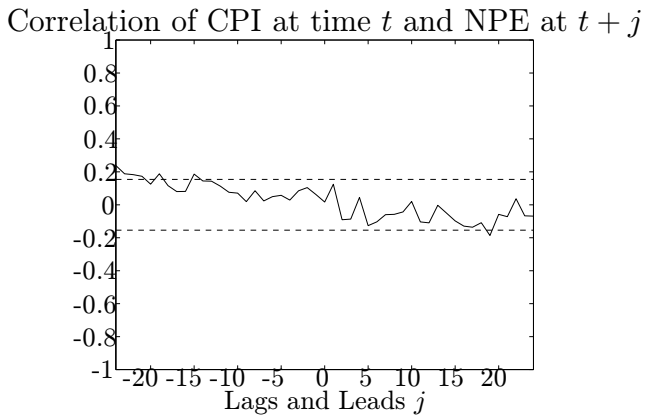


Figure 7: The cross correlation between CPI at time t and NPE at time $t+j$ together with approximate 95% confidence bounds ($\pm 2/\sqrt{T}$).

variable whether $c_0^i = 0$ and $c_1^i = 1$ when fitting

$$m^i(t) = c_0^i + c_1^i\, m_F^i(t) + \epsilon^i(t), \qquad i = CPI,\ NPE, \tag{14}$$

where ϵ^i is white noise. Unbiasedness cannot be rejected at the 1% level for NPE, but is strongly rejected for the CPI series for the period 1985-1998. Concentrating on the post-1994 subsample, that used for the yield-curve model, CPI forecasts also "pass" the unbiasedness test at the 1% confidence level.[22] Finally, three lagged values of CPI, NPE and the target rate were included on the right-hand side of (14), but none had a significant coefficient, except perhaps for a weak effect of the first lagged CPI on NPE. To conclude, analyst forecasts of CPI and NPE provide a reasonably good description of the conditional expected values of these variables, at least in the post-1994 period, so that (14) will be used with $c_0^i = 0$ and $c_1^i = 1$.

The second step in setting up the joint system for NPE and CPI is an examination of their correlation. The contemporaneous correlation between the two variables is small (0.017). When lagging one of these variables, the correlation estimates rarely exit approximate 95% confidence bounds around zero, as can be seen from Figure 7. Granger causation regressions (not reported here) in fact show that NPE does not help much in predicting future CPI. The CPI might, however, help in forecasting NPE. Including lagged values of the CPI in an AR(3) specification, for example, of NPE leads to a small gain in adjusted R^2 (9.5% to 12.1%). For each of CPI and NPE, we therefore specify conditionally independent subsystems, using the target rate as an exogenous variable.

The final requirements of the state-space system are that the macro variables are part of the state, and that the state is a four-dimensional autoregressive process of order 1 (this last choice will be further discussed below).

[22] Balduzzi, Elton, and Green (1997) and Li and Engle (1998) conduct this test with a short history as well, and fail to reject the null.

In the continuous-time economy, one deterministic counting process N^i records macroeconomic releases and another, N_F^i, counts the times at which analyst forecasts are made. Selecting a release time τ^i and the succeeding analyst forecast time τ_F^i, the specification can be summarized as:

$$\begin{aligned} m^i(\tau^i) &= m_F^i(\tau^i) + \epsilon^i(\tau^i) \\ m_F^i(\tau_F^i) &= a_0^i + a_1^i m_F^i(\tau_F^i-) + a_2^i m^i(\tau_F^i-) + a_3^i \theta(\tau_F^i-) + \epsilon_F^j(\tau_F^i), \end{aligned} \tag{15}$$

where $\epsilon^i(\tau^i)$ and $\epsilon_F^i(\tau_F^i)$ are jointly Gaussian independent across time with mean zero, respective variances σ^i and σ_F^i, and covariance σ_{mF}^i.[23]

Maximum-likelihood parameter estimates for (15) are reported in Table 7, except for the covariance parameter σ_{mF}^i, which was estimated to be zero for both CPI and NPE. The forecast m_F of m does not depend on the past value of m (both a_2^{CPI} and a_2^{NPE} are small and insignificant), and depend only slightly on the past target (a_3^{CPI} is somewhat larger than a_3^{NPE}, but neither is significant). The impulse reponse of a macro variable to a release surprise $\epsilon(\tau)$ at time τ thus dies off after one month, since the surprise does not affect future released values. In this sense, *release surprises are only temporary components of macro variables.* Based on (15), m^{CPI} and m^{NPE} are both treated as the sum of an AR(1), the forecast m_F, and Gaussian white noise. In other words, this specification results in an ARMA(1,1)-structure, which is also the specification selected by comparing Akaike and Schwarz criteria of ARMA(p,q) models that include p lagged values of the target rate. The restrictiveness of a four-dimensional state can be checked by including additional lagged values of m_F and m in the equation determining analyst forecasts. I do not find additional significant terms for the CPI forecasts, but an additional moving average for NPE forecasts improves the model-selection criteria.

6.3 Model-Implied Response to Release Surprises

The state vector X is now augmented by $(m(t), m_F(t))$, of which only $m(t)$ enters the stochastic intensities (2). If an FOMC meeting and a macro jump event happen on the same day, the Fed's target decisions are able to condition on the newly released information, as CPI and NPE releases (at 8:30 a.m.) precede FOMC meetings. The approximation of the likelihood function is derived in Appendix C.3. The estimation uses 1, 3, 6, and 12-month LIBOR rates, the target rate, CPI, NPE, and analyst forecasts. In addition to the weekly

[23]Fore more intuition, consider the problem of modeling m_F and m in discrete time. Let the state be denoted by $z = (z^{CPI}, z^{NPE})$ with $z^i \in R^2$, $i = CPI, NPE$. Leaving out the dependence on the target for the moment, the observation equations in the state space system are (1) $m_F^i(t+1) = \alpha_0 + \alpha_1 z_1^i(t) + \alpha_2 z_2^i(t)$ and (2) $m^i(t) = \alpha_0 + z_1^i(t)$. This shows that z_2 is the latent state that summarizes the information used by investors to form the forecast m_F. The state equation is $z^i(t) = A^i z^i(t-1) + u(t)$ with $u(t) \sim \mathrm{N}(0, \Omega^i)$, $A^i = \begin{pmatrix} 0 & 1 \\ \alpha_1^i & \alpha_2^i \end{pmatrix}$. This system is the maximally flexible system that imposes that (i) $m_F^i(t)$ is the conditional mean of $m^i(t)$, (ii) independence between CPI and NPE, (iii) m^i is part of the state z^i, and (iv) z^i is an AR(1). This is equivalent to (15).

Table 7: Joint Dynamics of Analyst Forecasts and Actual Releases

i	a_0^i	a_1^i	a_2^i	σ^i	σ_F^i
CPI	0.0061 (0.60)	0.5564 (3.08)	-0.0276 (-0.19)	0.0212 (7.30)	0.0182 (12.15)
NPE	0.0058 (1.08)	0.5493 (2.87)	-0.0117 (-0.11)	0.0168 (8.39)	0.0092 (10.16)

NOTE: This table reports maximum likelihood estimates of (15) using the sample 1985:2 to 1998:12 with t-ratios in brackets. The target rate $\theta(\tau_F^i)$ in (15) is the value of the target on the day before the analyst forecast survey τ_F^i, as releases occur at 8:30 am and FOMC meetings after that.

observations from Section 5, the day of and preceding CPI and NPE releases are included into the sample. Both daily data and longer maturities would raise the computational burden significantly, as was discussed in Section 4.4.

The impact of the surprise $m(\tau) - m_F(\tau)$ of a macroeconomic release at time τ on yields is determined by how it affects the arrival intensity of policy moves. The surprise can affect future Fed behavior in two ways. First, it could in principle change investors' forecast about future macro variables the Fed reacts to following λ_{CPI} and λ_{NPE} in (2). Surprises are, however, only temporary components of macro variables: they do not affect the future path of macro variables. Release surprises therefore have a 'short life', which means that they cannot affect macro variables that are relevant for FOMC meetings that are scheduled beyond the next macroeconomic release. Second, the surprise can affect the probability of a target change at the next FOMC meeting, which then changes future intensities according to λ_θ in (2). The estimation results (not reported here) show that this latter effect is not important, so that the estimated λ_{CPI} and λ_{NPE} are not significant.

This can also be seen from Figure 8 which plots the cross-sectional response of yields to a one-standard deviation release surprise to NPE. In the model, release surprises do not have hump-shaped coefficients like in the unrestricted regressions, but they are monotonically decreasing with maturity. This shows that release surprises are not 'inertia factors' like z. By introducing, for example, a jump in the inertia factor z at employment releases that is correlated with the actual surprise, the surprise may 'live longer,' as a shock to z affects intensities in the farer future.

The Effects of Employment Releases (in Data Points)

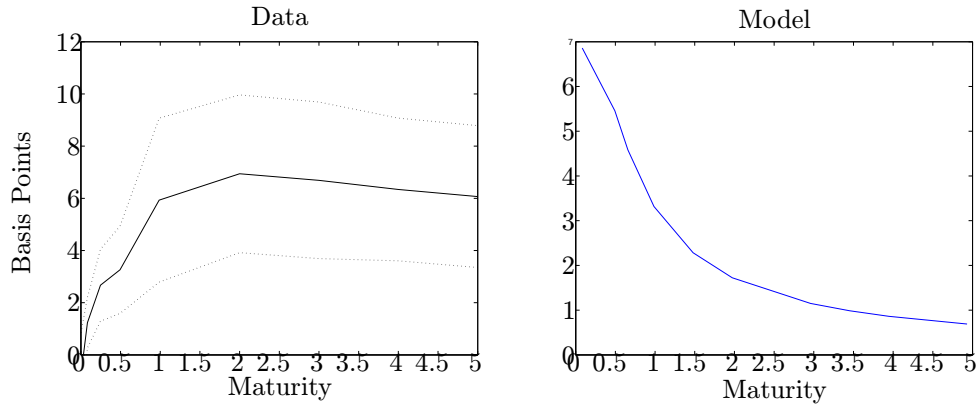


Figure 8: The first graph shows the slope parameter of daily yield changes regressed on NPE standardized surprises (and an intercept) using the subsample of the respective release days. Dotted lines are standard-error bounds computed with 5 Newey-West lags. Standardized surprises are defined as the analyst forecast error $m(t)-m_F(t)$, normalized by its standard deviation, so that the regression coefficients can be interpreted as reactions to a one-standard deviation analyst forecast error. The second graph shows the model-implied instantaneous response of yields to NPE release surprises. Due to time-zone differences, the data used for this estimation are same-day observations from 1994 to 1998 on 2, 3, 4, 5-year swap rates, and next-day observations of 1, 3, 6 and 12-month LIBOR rates.

References

Ahn, Dong-Hyun, Robert F. Dittmar and A. Ronald Gallant (2001), "Quadratic Term Structure Models: Theory and Evidence." Forthcoming in *Review of Financial Studies*.

Anderson, Torben G., Luca Benzoni and Jesper Lund (1998) "Estimating Jump-Diffusions for Equity Returns." Working Paper, J.L. Kellogg Graduate School of Management, Northwestern University.

Anderson, Torben G. and Jesper Lund (1996). "Stochastic Volatility and Mean Drift in the Short Rate Diffusion: Sources of Steepness, Level and Curvature in the Yield Curve." Working Paper, J.L. Kellogg Graduate School of Management, Northwestern University.

Babbs, Simon H. and Nick J. Webber (1993). "A Theory of the Term Structure with an Official Interest Rate." Working Paper, Warwick Business School, Warwick University.

Babbs, Simon H. and Nick J. Webber (1996). "Term Structure Modeling under Alternative Official Regimes" in "Mathematics of Derivative Securities," eds Dempster and Pliska, Cambridge University Press.

Balduzzi, Pierluigi, Giuseppe Bertola and Silverio Foresi (1996). "A Model of Target Changes and the Term Structure of Interest Rates." *Journal of Monetary Economics* vol. 39, no. 2, 223-49.

Balduzzi, Pierluigi, Giuseppe Bertola, Silverio Foresi and Laura Klapper (1998). "Interest Rate Targeting and the Dynamics of Short-Term Rates." *Journal of Money, Credit and Banking* 30:1, 26-50.

Balduzzi, Pierluigi, Edwin J. Elton and T. Clifton Green (1998). "Economic News and the Yield Curve: Evidence from the U.S. Treasury Market." Working Paper, Boston College.

Berardi, Andrea (1998). "Term Structure, non-neutral inflation and economic growth: a three factor model." Working Paper, London Business School.

Brandt, Michael and Pedro Santa-Clara (2001) "Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets." Forthcoming in *Journal of Financial Economics.*

Brémaud, Pierre (1981). "*Point processes and queues, martingale dynamics.*" Berlin: Springer-Verlag.

Buraschi, Andrea and Alexei Jiltsov (2000). "How Large is the Inflation Risk Premium? A Monetary Model of the Yield Curve." Working Paper, London Business School.

Campbell, John Y. and Luis M. Viceira (2001). "Who Should Buy Long Term Bonds?" Forthcoming in *American Economic Review.*

Campbell, John Y. and Robert J. Shiller (1991). "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Review of Economic Studies*, May, 495-514.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (1998). "Monetary Policy Shocks: What Have We Learned and to What End?" prepared for the *Handbook of Macroeconomics.*

Cochrane, John (1998). "Where is the Market Going? Uncertain Facts and Novel Theories." *NBER Working Paper* no 6207.

Collin-Dufresne, Pierre and Bruno Solnik (2001). "On the term structure of default premia in the swap and LIBOR markets." Forthcoming in *Journal of Finance.*

Constantinides, George (1992). "A Theory of the Nominal Term Structure of Interest Rates." *Review of Financial Studies* v5, n4, 531-52.

Cox, John C., Jonathan E. Ingersoll and Stephen A. Ross (1985). "A Theory of the Term Structure of Interest Rates." *Econometrica*, vol 53, no.2, 385-407.

Dai, Qiang and Kenneth J. Singleton (2000). "Specification Analysis of Affine Term Structure Models." *Journal of Finance*, vol. LV, 1943-78.

Das, Sanjiv R. (1001). "The Surprise Element: Jumps in Interest Rates." Forthcoming in *Journal of Econometrics*.

Delbaen, Freddy and W. Schachermayer (1994). "A General Version of the Fundamental Theorem of Asset Pricing." *Math. Ann.*, vol 300, 463-520.

Duffie, Darrell and Jun Liu (2001). "Floating-Fixed Credit Spreads." Forthcoming in *Financial Analyst Journal*.

Duffie, Darrell and Rui Kan (1996). "A Yield-Factor Model of Interest Rates." *Mathematical Finance*, vol. 6 (1996), no. 4, 379-406.

Duffie, Darrell, Jun Pan, and Kenneth Singleton (2000). "Transform Analysis and Asset Pricing for Affine Jump-Diffusions." *Econometrica*, Vol. 68, 1343-1376.

Duffie, Darrell and Kenneth Singleton (1997). "An Econometric Model of the Term Structure of Interest Rate Swap Yields" *Journal of Finance* 52, pp. 1287-1323.

Eichenbaum, Martin and Charles E. Evans (1995). "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates." *Quarterly Journal of Economics* 110, 975-1009.

El Karoui, Nicole, R. Myneni and R. Viswanathan (1993). "Arbitrage Pricing and hedging of interest rate claims with state variables." Working Paper, Université de Paris VI, Laboratoire de Probabilité'.

Evans, Chales L. and David Marshall (1998). "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory." *Carnegie-Rochester Conference Series on Public Policy*, Vol. 49, 53-111.

Farnsworth, Heber and Richard Bass (2000). "The Term Structure with Credible and Semi-Credible Targeting." Working Paper, Washington University in St. Louis.

Fleming, J. Michael and Eli M. Remolona (1997). "What Moves the Bond Market ?" *FRBNY Economic Policy Review* December 1997, 31-50.

Fleming, J. Michael and Eli M. Remolona (1999). "The Term Structure of Announcement Effects." Working Paper, Federal Reserve Bank of New York.

Gallant, A. Ronald and George Tauchen (1996). "Which Moments to Match?" *Econometric Theory* 12, 627-681.

Geweke, John (1996). "Monte Carlo Simulation and Numerical Integration" in *Handbook of Computational Economics* Volume 13, Amman, Hans M.Kendrick, David A.Rust, John, eds., Elsevier Science, North-Holland, Amsterdam.

Hamilton, James D. (1996). "The Daily Market for Federal Funds." *Journal of Political Economy*, vol. 104, no1., pp. 26-56.

Hamilton, James and Oscar Jorda (2000). "The Stance of Monetary Policy: The Federal Funds Rate Target? Searching in a New Direction." Working Paper, San Diego.

Hansen, Lars P. and Ravi Jagannathan (1991). "Implications of Security Market Data for Models of Dynamic Economies." *Journal of Political Economy* v99, n2, pp. 225-62.

Harrison, Michael and David Kreps (1979). "Martingales and Arbitrage in Multiperiod Security Markets." *Journal of Economic Theory* 20: 381-408.

Harrison, Michael and Stanley Pliska (1981). "Martingales and Stochastic Integrals in the Theory of Continuous Trading." *Stochastic Processes and Their Applications* 11: 215-260.

Honoré, Peter (1997). "Modeling Interest Rate Dynamics in a Corridor with Jump Processes." Working Paper, Aarhus School of Business.

Johannes, Michael (2001). "A Nonparametric Jump-Diffusion Model of the Short Rate." Working Paper, Columbia University.

Jones, Charles M., Owen Lamont, and Robin Lumsdaine (1996). "Macroeconomic News and Bond Market Volatility." *Journal of Financial Economics* 47, 315-337.

Li, Li and Robert F. Engle (1998). "Macroeconomic Announcements and Volatility of Treasury Futures." Working Paper No. 98-27, UC San Diego.

Litterman, Robert and José Scheinkman (1993). "Common factors affecting bond returns." *Journal of Fixed Income* 1, 51-61.

Litterman, Robert, José Scheinkman and Laurence Weiss (1988). "Volatility and the Yield Curve." Working Paper, Goldman & Sachs, Financial Strategies Group.

Lo, Andrew W. (1988). "Maximum Likelihood Estimation of Generalized Ito Processes with Discretely-Sampled Data." *Econometric Theory* 4, pp. 231-247.

Konstantinov, Vassil (2000). "Fed Funds Rate Targeting, Monetary Regimes and the Term Structure of Interbank Rates." Working Paper, University of Wyoming.

Merton, Robert (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* v. 29, pp. 449-470.

Meulendyke, Ann-Marie (1998). "*U.S. Monetary Policy & Financial Markets.*" Federal Reserve of New York.

Orphanides, Athanasios (1998). "Monetary Policy Rules Based on Real-Time Data." Working Paper, Finance and Economics Discussion Series, Federal Reserve Board Washington.

Pearson Neil D. and Tong-Sheng Sun (1994). "Exploiting the Conditional Density in estimating the Term Structure: An application to the Cox, Ingersoll and Ross model." *Journal of Finance* XLIX(4): 1279-1304.

Pedersen, Asger Roer (1995). "A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations." *Scandinavian Journal of Statistics* 22, 55-71.

Protter, Philip (1990). "*Stochastic Integration and Differential Equations.*" Berlin: Springer Verlag.

Rudebusch, Glenn D. (1995). " Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure." *Journal of Monetary Economics*, v35, n2, 245-74.

Sack, Brian (1998). "Does the Fed Act Gradually? A VAR Analysis." Working Paper, Board of Governors of the Federal Reserve System.

Singleton, Kenneth (2001). "Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function." Forthcoming in *Journal of Econometrics.*

Taylor, John B. (1993). "Discretion versus policy rules in practice." *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 195-214, 1993.

Vasicek, Oldrich A. (1977). "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics* v5, n2, pp. 177-88.

Woodford, Michael (1999). "Optimal Monetary Policy Inertia." *NBER Working Paper* No 7261, July 1999.

Appendices

A Linear-Quadratic Jump-Diffusions

Fix a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and a filtration $\{\mathcal{F}(t) : t \geq 0\}$ satisfying the usual conditions (Protter (1990)). Linear-quadratic jump-diffusions (LQJDs) are defined by choosing particular functional forms for the coefficients μ and σ of the SDE

$$dX(t) = \mu(X(t), t)\, dt + \sigma(X(t, t))\, dW(t) + dJ(t),$$

where W is a vector of Brownian motions, together with additional restrictions on the jump process J. In describing these parametric specifications, I can, without loss of generality, partition the state as $X = (X_1, X_2)$ so that X_2 is a k_2-dimensional process, with $k_1 + k_2 = N$. Assumption 1 will restrict X_2 to be Gauss-Markov. It will be convenient to define the set

$$C = \{(c_0, c_1, c_2) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^{N\times N} : c_2 \text{ is symmetric positive semidefinite} \\ \text{and consists of zeros except possibly the lower right } k_2 \times k_2 \text{ partition}\}$$

of coefficients. I will make repeated use of LQ functions of the state of the form $g : D \times C \to \mathbb{R}$, with

$$g(x, c) = c_0 + c_1\, x + x^\top c_2 x. \tag{A.1}$$

I am now in the position to specify the LQJD as follows.

Assumption 1 (Characterization of LQJD processes)

(a) (Functional Form of Drift and Volatility)

The drift and 'volatility' of X are given by are given by

$$\mu(x, t) = K(t)\ (\bar{x}(t) - x) \tag{A.2}$$

$$\sigma(x, t) = \Sigma(t)\, S(x, t), \tag{A.3}$$

where $S(x, t)$ is a $N \times N$ diagonal matrix with i-th diagonal element $[S(x, t)]_{i,i} = \sqrt{s_{0i}(t) + s_{1i}(t) \cdot x}$, and where the coefficients $s_{0i}(t) \in \mathbb{R}$, $s_{1i}(t), \bar{x} \in \mathbb{R}^N$ and $K(t), \Sigma(t) \in \mathbb{R}^{N\times N}$ are deterministic functions of time.

(b) (Functional Form of Stochastic Intensities)

The jumps J are counted by a p-dimensional counting process N_p with stochastic intensity, and by a d-dimensional deterministic counting process N_d with no common jump times.[24] The stochastic intensity $\{\lambda_i(t) : t \geq 0\}$ of N_p^i is given by

$$\lambda_i(t) = g\left(X(t-), l^i(t)\right), \tag{A.4}$$

for time-dependent coefficients $l^i(t) \in C$. The coefficients $l^i(t)$ and the domain D satisfy joint conditions to ensure that $\lambda_i(t) \geq 0$, as required for any intensity process.

[24] This means that $\Delta N_p^i \cdot \Delta N_p^j = 0$ and $\Delta N_d^i \cdot \Delta N_d^j = 0$, $i \neq j$ almost surely.

(c) (Conditional Jump-Size Distributions)

For any Poisson jump time τ, the $\mathcal{F}(\tau-)$-conditional distribution $v_{p,\tau}$ of the jump size $\Delta X(\tau)$ is independent of $X(\tau-)$. For any deterministic jump time t, the $\mathcal{F}(t-)$-conditional distribution $v_{d,t}$ of the jump size $\Delta X(t)$ has a Laplace transform which is an exponential LQ function of $X(t-)$. More precisely, for all $a \in \mathbb{R}^N$, I have that

$$E_{t-}\left[\exp(a \cdot \Delta X(t))\right] = \exp\left(g\left(X(t-), c(t;a)\right)\right) \tag{A.5}$$

for some $c(t;a) \in C$.

(d) (Parameter Restrictions)

(i) All of the time-dependent coefficients are bounded and piece-wise constant functions of time.[25]

(ii) Joint restrictions on (μ, σ, J) and the domain D apply that guarantee a unique (strong) solution to (4).

(iii) Gaussianity of X_2: The lower left $k_2 \times k_1$ partitions of the matrices $K(t)$ and $\Sigma(t)$, labeled $K_{21}(t)$ and $\Sigma_{21}(t)$, consist of zeros only. Also, $s_{1i}(t)$ is an N-vector of zeros for all $i \in \{k_1+1, \ldots, N\}$. Finally, the lower k_2 entries of J are always zero (X_2 does not jump).

B Linear-Quadratic Yield-Curve Model

In the specification estimated in this paper, the short rate r is taken to be the sum of two state variables, the spread and the target. More generally, the affine structure of Duffie and Kan (1996) and the quadratic structure of the SAINTS model Constantinides (1992), El Karoui, Myneni, and Viswanathan (1993) can be combined by the following assumption.

Assumption 2 (Linear-Quadratic Short Rate of Interest)

Fixing a linear-quadratic jump diffusion X, the short-rate process $\{r(t);\ t \geq 0\}$ is assumed to have the linear-quadratic form $R(x,t) = g(x, \delta(t))$ for some given coefficients $\delta(t) \in C$.

B.1 Statement of Lemma 1

Let the times at which deterministic jumps occur between t and T be denoted $\tau_1^d, \ldots, \tau_n^d$. In the following, Assumption 1 will often be assumed to hold under the risk-neutral probability measure $\mathcal{Q}$. This will be taken to mean that (A.2), (A.3) and (A.4) define the coefficients

[25] This particular type of time-dependence of the parameters determining the dynamics of X is sufficient for the seasonality effects studied this paper. Alternatively, the parameters may be bounded continuous functions of time.

of the dynamics of the state (4) under $\mathcal{Q}$, and (A.5) will hold with the expectation taken under $\mathcal{Q}$.

Lemma 1: Suppose that Assumptions 1 and 2 hold under $\mathcal{Q}$. Additionally, for $t = \tau_i^d$, for some i, suppose that $P(t,T) = \exp\left(g\left(X(t), \bar{c}\,\right)\right)$ and for some $\bar{c} = (\bar{c}_0, \bar{c}_1, \bar{c}_2) \in C$. Then there exist coefficients $c \in C$ such that $P(t-,T) = \lim_{s\uparrow t} P(s,T)$ is given by

$$P(t-,T) = \exp\left(g\left(X(t-), c\,\right)\right). \tag{B.6}$$

Proof: From equation (6),

$$\begin{aligned}
P(t-,T) &= E_{t-}^{\mathcal{Q}}\left[P\left(t,T\right)\right] \\
&= E_{t-}^{\mathcal{Q}}\left[\exp\left(g\left(X(t), \bar{c}\,\right)\right)\right] \\
&= \exp\left(g\left(X(t-), \bar{c}\,\right)\right) E_{t-}^{\mathcal{Q}}\left[\exp\left(\bar{c}_1 \cdot \Delta X(t)\right)\right] \\
&= \exp\left(g\left(X(t-), \bar{c}\,\right)\right) \exp\left(g\left(X(t-), a(t;\bar{c}_1)\right)\right) \\
&= \exp\left(g\left(X(t-), \bar{c} + a(t;\bar{c}_1)\right)\right),
\end{aligned}$$

where $E_{t-}^{\mathcal{Q}}$ denotes $\mathcal{F}(t-)$-conditional expectation under $\mathcal{Q}$, and the fourth equality holds for some $a(t;\bar{c}_1) \in C$ because of Definition (A1.c.). ■

B.2 Statement of Lemma 2

Lemma 2: Suppose that, for some i such that $s = \tau_{i+1}^d$, $P(s-,T) = \lim_{t\uparrow s} P(t,T)$ can be represented as $P(s-,T) = \exp\left(g\left(X(s-), \bar{c}\,\right)\right)$ for some $\bar{c} \in C$. Let Assumptions 1 and 2 be satisfied under $\mathcal{Q}$. Also suppose that the integrability conditions hold at $(s, \bar{c})$, as defined below. Then for each $t \in [\tau_i^d, s)$, there exist coefficients $c(t,s) \in C$ such that

$$P(t,T) = \exp\left(g\left(X(t), c(t,s)\,\right)\right), \tag{B.7}$$

where $c(\,\cdot\,, s) : [\tau_i^d, s] \to C$ solves the system of ordinary differential equations (ODE's) in (B.11)-(B.13) stated below, with terminal condition $c(s,s) = \bar{c}$.

Proof: Lemma 2 applies the standard Feynman-Kac approach to equation (6) between deterministic jump times. The approach proceeds in two steps. Step 1 states and solves the relevant Cauchy problem. Step 2 imposes integrability conditions so that the bond price at time $t \in [\tau_i^d, s)$ can indeed be viewed as the Feynman-Kac solution to the Cauchy problem of Step 1.

Step 1 of Proof: Consider the following Cauchy problem. For all $t \in [\tau_i^d, s)$ and

$x \in D$, let $F(t,s,x)$ solve the partial differential-integral equation (PDIE)

$$\begin{aligned} 0 &= F_t(t,s,x) + F_x(t,s,x) \cdot \mu(x,t) \\ &+ \frac{1}{2}\, tr\, \left[F_{x\,x}(t,s,x)\sigma(x,t)\sigma(x,t)^\top\right] \\ &+ \sum_{i=1}^{p} g(x,l_i(t)) E^{\mathcal{Q}}\left[F(t,s,x+J_i^p(t)) - F(t,s,x)\right] - R(x,t)F(t,s,x), \end{aligned} \tag{B.8}$$

with terminal condition $F(s,s,x) = \exp\left(g\left(x,\bar{c}\right)\right)$. Here, the notation "$E^{\mathcal{Q}}\left[J_i^p(t)\right]$" stands for the unconditional mean of the distribution of the jump $\Delta X(t)$ in the state at an arbitrary jump time τ_i^p of N_i^p. Because this jump is of a distribution independent of $X(t-)$, and because the number of jumps during any time interval is finite almost surely, the expectation is unambigeous, despite the abuse of notation. Guess a solution of the form

$$F(t,s,x) = \exp\left(g\left(x, c(t,s)\right)\right), \tag{B.9}$$

where the coefficients $c(t,s) = (c_0(t,s), c_1(t,s), c_2(t,s))$ satisfy terminal conditions at s given by $\bar{c} = (\bar{c}_0, \bar{c}_1, \bar{c}_2)$. Now I verify that the guess in (B.9) solves the PDIE (B.8) for all $t \in [t_i, s)$. By applying Ito's Lemma to (B.9) and using the fact that $F(t,s,x)$ is strictly positive, I have

$$\begin{aligned} 0 &= \frac{dc_0(t,s)}{dt} + \frac{dc_1(t,s)}{dt} \cdot x + x^\top \frac{dc_2(t,s)}{dt}\, x \\ &+ \left(c_1(t,s) + 2\, c_2(t,s)\, x\right) \cdot K(t)(\bar{x}(t) - x) \\ &+ \frac{1}{2} tr[\, \big((c_1(t,s) + 2c_2(t,s)x)(c_1(t,s) + 2c_2(t,s)x)^\top + 2c_2(t,s)\big) \\ &\big(\Sigma(t)\, S(x,t) S(x,t)^\top \Sigma(t)^\top\big)] \\ &+ \sum_{i=1}^{p} g\left(x, l_i(t)\right) E^{\mathcal{Q}}\left[\exp(c_1(t,s) \cdot J_p^i(t)) - 1\right] - g\left(x, \delta(t)\right), \end{aligned} \tag{B.10}$$

where the coefficients with subscripts are subvectors and submatrices of the coefficients in equations (A.2), (A.3), (A.4), and (A.5). This equation must hold for all $x \in D$, which is assumed to contain an open set, so that I can apply the usual method of undetermined coefficients which equates the coefficients of x and the quadratic forms in x to zero. This shows that $c(t,s)$ solves the ODE's:

$$
\begin{aligned}
\frac{dc_0(t,s)}{dt} &= \delta_0(t) - c_1(t,s)^\top K(t)\,\bar{x}(t) \\
&- \frac{1}{2}\sum_{i=1}^{N}[c_1(t,s)^\top \Sigma(t)]_i^2\, s_{i0} \\
&- \frac{1}{2} tr[2\, c_3(t,s)\Sigma(t)S(x,t)S(x,t)^\top \Sigma(t)^\top] \\
&- \sum_{i=1}^{p} l_{0,i}(t)\; E^{\mathcal{Q}}\left[\exp(c_1(t,s)\cdot J_p^i(t)) - 1\right]
\end{aligned} \tag{B.11}
$$

$$
\begin{aligned}
\frac{dc_1(t,s)}{dt} &= \delta_1(t) + K(t)^\top c_1(t,s) - 2\, c_2(t,s)K(t)\,\bar{x}(t) \\
&- \frac{1}{2}\sum_{i=1}^{N}[c_1(t,s)^\top \Sigma(t)]_i^2 s_{i1}(t) \\
&- 2\, c_2(t,s)\,\Sigma(t)S(x,t)S(x,t)^\top \Sigma(t)^\top c_1(t,T) \\
&- \sum_{i=1}^{p} l_{1,i}(t)\; E^{\mathcal{Q}}\left[\exp(c_1(t,s)\cdot J_p^i(t)) - 1\right]
\end{aligned} \tag{B.12}
$$

$$
\begin{aligned}
\frac{dc_2(t,s)}{dt} &= \delta_2(t) - c_2(t,s)K(t) - K(t)^\top c_2(t,s) \\
&- 2\, c_2(t,s)\,\Sigma(t)S(x,t)S(x,t)^\top \Sigma(t)^\top\, c_2(t,s) \\
&- \sum_{i=1}^{p} l_{2,i}(t) E^{\mathcal{Q}}\left[\exp(c_1(t,s)\cdot J_p^i(t)) - 1\right],
\end{aligned} \tag{B.13}
$$

with terminal conditions given by $\bar{c}_0, \bar{c}_1$, and $\bar{c}_2$, respectively. ■

Step 2 of Proof: The *integrability conditions hold at* $(s, \bar{c})$ if

1. $c(\cdot, s) : [\tau_i^d, s] \to C$ uniquely solve equations (B.11)-(B.13) with terminal conditions $\bar{c}$ at time s.

2. $E^{\mathcal{Q}} \int_0^s |\gamma_1^i| \, dt < \infty$, for all $i = 1, \ldots, p$, where
$\gamma_1^i(t) = \Psi(t-) \, E^{\mathcal{Q}} \left[\exp(c_1(t,s) \cdot J_p^i(t)) - 1\right] g\left(X(t-), l_i(t)\right).$

3. $E^{\mathcal{Q}} \left(\int_0^s |\gamma_2(t) \cdot \gamma_2(t)| \, dt\right)^{1/2} < \infty$, where
$\gamma_2(t) = \Psi(t-) \left[c_1(t,s) + 2\, X(t-)^\top c_2(t,s)\right] \sigma(X(t-), t).$

4. $E^{\mathcal{Q}} \left(|\Psi(s)|\right) < \infty$,

where $\Psi(t)$ is defined for $t \in [\tau_i^d, s]$ by

$$\Psi(t) = \begin{cases} \exp\left(-\int_0^t R(X(u), u) du\right) \exp\left(g\left(X(t), c(t,s)\right)\right), & \text{for} \quad t \in [\tau_i^d, s) \\ \exp\left(-\int_0^s R(X(u), u) du\right) \exp\left(g\left(X(s-), \bar{c}\right)\right), & \text{for } t = s. \end{cases} \tag{B.14}$$

If the integrability conditions hold at $(s, \bar{c})$, then $\Psi(t)$ given by (B.14) is a martingale for $t \in [\tau_i^d, s]$. This can be seen by applying Ito's Lemma[26] to equation (B.14) for $t \in [\tau_i^d, s]$ and using the coefficient calculation (B.11)-(B.13) gives

$$\begin{aligned} d\Psi(t) = \Psi(t-) \left[\widehat{c}_1(t,s) + 2\, X(t-)^\top \widehat{c}_2(t,s)\right] \sigma(X(t-), t) \; dW(t) + \\ + \sum_{i=0}^{p} \Psi(t-) \; \left[\exp\left(\widehat{c}_1(t,s) \cdot J_p^i(t)\right) - 1\right] \; dM_P^i(t), \end{aligned}$$

where M_P^i denotes the i-th compensated Poisson process. Duffie, Pan, and Singleton (2000), p. 26, show that with assumptions 4.1 and 4.2, $\int \eta_2(t) dW$ and $\int \Psi(t-) \; \left[\exp\left(\widehat{c}_1(t,s) \cdot J_p^i(t)\right) - 1\right] \; dM_P^i$, for $i = 1, \ldots, p$, are martingales during the interval $[\tau_i^d, s]$.

This implies that, for $t \in [\tau_i^d, s)$,

$$P(t,T) = E_t^{\mathcal{Q}} \left[\exp\left(-\int_t^s R(X(u), u) du\right) \; \exp\left(g\left(X(s-), \bar{c}\right)\right)\right]$$

can be viewed as the Feynmac-Kac solution to the Cauchy problem (B.8).

[26] See Protter (1990), p. 74.

B.3 Statement of Proposition 1

Proposition 1: Suppose that Assumptions 1 and 2 hold under $\mathcal{Q}$. Let the coefficient vector $c\,(t,T)$ be calculated recursively using the algorithm shown in Appendix B.4. Assume that the integrability conditions hold at $(\tau_i^d, c(\tau_i^d, T))$, $i \in \{1, \ldots, n\}$, and also at $(T, 0)$. Then $P(t,T) = \psi\,(X\,(t)\,, c(t,T))$ for all $t \leq T$.

Proof: The proof is by induction over the deterministic jump dates $\tau_1^d, \ldots, \tau_n^d$ between t and T. By assumption, $P(T,T) = 1$. Applying Lemma 2 with $s = T$, $\bar{c} = 0$ and $\psi(X(T), \bar{c}) = 1$, implies that $P(t,T)$ satisfies (7) for $t \in [\tau_n^d, T)$. Lemma 1 can then be applied to obtain the desired property for $P(\tau_n^d-, T)$. Now suppose, for any deterministic jump time τ_i^d, that $P(\tau_{i+1}^d-, T)$ is given by (7). Lemma 2 can be applied to establish the desired property for any time $t \in [\tau_i^d, \tau_{i+1}^d)$ and then Lemma 1 to get it for $P(\tau_i^d-, T)$. By induction, $P(t,T)$, $t \in [0,T]$, has the desired property. (Note that Lemma 2 can also be applied to the interval $[0, \tau_1^d)$).■

B.4 Recursive Calculation of $c(t,T)$

The algorithm for computing $c(t,T)$ is provided here.

Step 0 (Initialization):

The terminal condition for $c(t,T)$ at T consists of a collection of zeros denoted by $\bar{c}_{n+1}$ in C. Let $\tilde{c}_n(t,T)$ solve the ODE's in (B.11)-(B.13) during the interval $[\tau_n^d, T]$ with terminal condition $\bar{c}_{n+1}$, and define $\tau_{n+1}^d = T$.
Go to Step 1.

Step i, for $i = 1, \ldots, n$:

- Calculate the new terminal condition for time τ_{n+1-i}^d as

$$\bar{c}_{n+1-i} = \tilde{c}_{n+1-i}\left(\tau_{n+1-i}^d, \tau_{n+2-i}^d\right) + c_{n+1-i}\left(\tau_{n+1-i}^d, \tilde{c}_{n+1-i}(\tau_{n+1-i}^d, \tau_{n+2-i}^d)\right),$$

  where $c_{n+1-i}\left(\tau_{n+1-i}^d, \tilde{c}_{n+1-i}(\tau_{n+1-i}^d, \tau_{n+2-i}^d)\right) \in C$ is taken from equation (A.5) evaluated at $t = \tau_{n+1-i}^d$.

- For a given terminal condition $\bar{c}_{n+1-i} \in C$, let $\tilde{c}_{n-i}(t,T)$ solve the ODE's in (B.11)-(B.13) during the interval $[\tau_{n-i}^d, \tau_{n+1-i}^d]$, with terminal condition $\bar{c}_{n+1-i}$.

Stop if $i = n$. Go to Step $i+1$.

Coefficient Collection:

The coefficients $c(t,T)$ are then equal to $\tilde{c}_i(t,T)$ for any $t \in (\tau_i^d, \tau_{i+1}^d)$ and equal to $\bar{c}_i$ at any $t = \tau_i^d$.

C Simulated Maximum Likelihood with Jumps

C.1 SML without Macro Variables

Suppose X contains the target rate θ, modeled as the (observable) difference of the "up" and "down" counting processes, with state-and time-dependent intensities as in (1). I abstract for the moment from the time dependence of stochastic intensities introduced by FOMC meetings, assuming that these intensities are always "active." Starting from $\tilde{x}$ at time $\tilde{t}$, X can be simulated with the scheme

$$\begin{aligned} \Delta \hat{X}_t^{\tilde{x}} &= \mu(\hat{X}_{t-h}^{\tilde{x}}, t-h)\, h + \sqrt{h}\, \sigma(\hat{X}_{t-h}^{\tilde{x}}, t-h)\, \epsilon_t + J_t^X z_t \\ \hat{X}_{\tilde{t}}^{\tilde{x}} &= \tilde{x}, \end{aligned} \tag{C.15}$$

where ϵ_t is *i.i.d.* standard normal and J^X is the deterministic jump[27] in X at random times, determined by a 2-dimensional vector of Bernoulli variables z_t that determine jumps, up and down. Using the conditional independence of the counting processes N^U and N^D, and assuming that the econometrician observes only the difference of the two, the simulation rolls a "three-sided die" to determine z_t. The three sides are "up" ("U", meaning $\theta_t - \theta_{t-h} = J_\theta$), "down" ("D", meaning $\theta_t - \theta_{t-h} = -J_\theta$), and "no change" ("0", meaning $\theta_t = \theta_{t-h}$). Their conditional probabilities at time t are approximately

$$p_{h,t}^j = \begin{cases} \lambda_{t-h}^U h\,(1 - \lambda_{t-h}^D h), & \text{for} \quad j = \text{U}, \\ \lambda_{t-h}^D h\,(1 - \lambda_{t-h}^U h), & \text{for} \quad j = \text{D}, \end{cases}$$

and $p_{h,t}^0 = p_{h,t}^U p_{h,t}^D + (1 - p_{h,t}^U)(1 - p_{h,t}^D)$.

Let X^θ denote all variables in X other than the target θ. The Monte-Carlo approximation of the conditional density is

$$f_X(X(t), t \,|\, \tilde{x}, \tilde{t}) \approx \frac{1}{S} \sum_{s=1}^{S} \sum_{i=\{U,D,0\}} \phi(X_t^\theta, t \,|\, \theta_t, \hat{X}_{t-h}^{\tilde{x}}[s], t-h)\, \hat{p}_{h,t}^i[s] 1_{i,t}[s], \tag{C.16}$$

where $\phi(\,\cdot\,, t \,|\, \hat{X}_{t-h}^{\tilde{x}}[s], t-h)$ is the Gaussian density of X_t at time t conditional on the value $\hat{X}_{t-h}^{\tilde{x}}[s]$ at time $t-h$, $\hat{X}_{t-h}^{\tilde{x}}[s]$ denotes the s-th simulated path from the scheme (C.15),

[27] The notation goes through with Gaussian jumps J^X, but needs to be adjusted in the case of other jumps size distributions.

$1_{i,t}[s]$ is the indicator for the i-th side of the die at time t in the s-th simulation, and $\hat{p}^i_{h,t}[s]$ is constructed using $\hat{X}^{\tilde{x}}_{t-h}[s]$. Let $\hat{\theta}^{\tilde{x}}_{t-h}$ be the target component of $\hat{X}^{\tilde{x}}_{t-h}$. If the simulated target $\hat{\theta}^{\tilde{x}}_{t-h}$ at time $t-h$ cannot reach the observed time t-value of target in at most one jump, that simulation is assigned zero likelihood.

I now turn to a case with *time-dependent* intensities that is relevant in Section 3. In that application, policy interventions on meeting days are modeled by activating state-dependent Poisson intensities only during FOMC meeting days. More precisely, suppose the i-th meeting day is during the interval $[\tilde{t}_M(i), t_M(i)]$. It is straightforward to modify the simulation scheme (C.15) to allow jumps only during such meeting-day intervals. The s-th path drawn from this modified scheme is referred to in what follows as $\hat{X}^{\tilde{x}}_t[s]$. I now construct analogues of (C.16) for this time-dependent case. As long as the observation time t lies within a meeting-day interval, in that $\tilde{t}_M(i) \leq t < t_M(i)$, the approximation (C.16) itself still applies. If the observation time t is made outside an FOMC meeting, however, then one might want to replace the Bernoulli-density terms with an indicator function for sample paths leading up to the actual value of the target at t,

$$f_X(X_t, t \,|\, \tilde{x}, \tilde{t}) \approx \frac{1}{S} \sum_{s=1}^{S} \phi\left(X^\theta_t, t | \hat{X}^{\tilde{x}}_{t-h}[s], t-h\right) 1_{\theta_t = \hat{\theta}^{\tilde{x}}_{t-h}[s]}. \tag{C.17}$$

In (C.17), jumps in the target enter the SML objective function only through the indicator function and the simulated values $\hat{X}^{\tilde{x}}_{t-h}$. This creates a serious problem when maximizing the objective: For a given (finite) number S of simulations, a small change in the parameter vector does not necessarily affect the average number of jumps across simulations and may thus leave the value of the likelihood function unchanged. Only changes in a parameter that are large enough to affect the number of simulated jumps change the objective function, but possibly by a large amount.[28] In order to overcome this discontinuity, an alternative to (C.17) is constructed as follows. The joint conditional density of factors can be written in the form

$$f_X\left(X_t, t \,|\, \tilde{x}, \tilde{t}\right) = f_\theta\left(\theta_t, t \,|\, \tilde{x}, \tilde{t}\right) \, f_{X^\theta|\theta}\left(X^\theta_t, t \,|\, \theta_t, \tilde{x}, \tilde{t}\right). \tag{C.18}$$

[28] A similar issue is encountered by Anderson, Benzoni, and Lund (1999) who estimate a jump-diffusion model for equity returns with EMM. In their specification, the size of the jump is Gaussian and the occurrence of jumps is not observed, so that they smooth the mapping from parameters to the estimation's objective function by allowing for partial jumps. In the setup considered in this paper, the target only moves in observable 25 basis points increments, so that a simulation of partial jumps is not feasible. A conjecture is that the method proposed in the following can also be applied with EMM, since the efficiency results of Gallant and Tauchen (1996) do not rely on a particular leading density.

The first term of equation (C.18) can be approximated by

$$\begin{aligned} f_\theta\left(\theta_t, t \,|\, \tilde{x}, \tilde{t}\right) &\approx \frac{1}{S}\sum_{s=1}^{S} f_\theta\left(\theta_t, t \,|\, \hat{X}^{\tilde{x}}_{t_M(i)-h}[s], t_M(i)-h\right) \\ &\approx \frac{1}{S}\sum_{s=1}^{S}\sum_{i=U,D,0} \hat{p}^{i}_{h,t_M(i)}[s] 1_{i,t}[s] \\ &\approx \bar{S}/S, \end{aligned}$$

where $\bar{S}$ denotes the total number of simulated paths that resulted in the observed value $\theta(t)$. In words, $\bar{S}/S$ is the frequency of "correctly simulated" target rates in the simulations (starting with $\tilde{x} = X_{\tilde{t}}$), while the expression in the first row weights the simulated paths by their likelihoods. In practice, with time-dependent intensities h must be chosen carefully, as intensities can become large during an FOMC meeting. Details about the choice of h can be found in Appendix C.2.

The second term in (C.18) can be approximated by

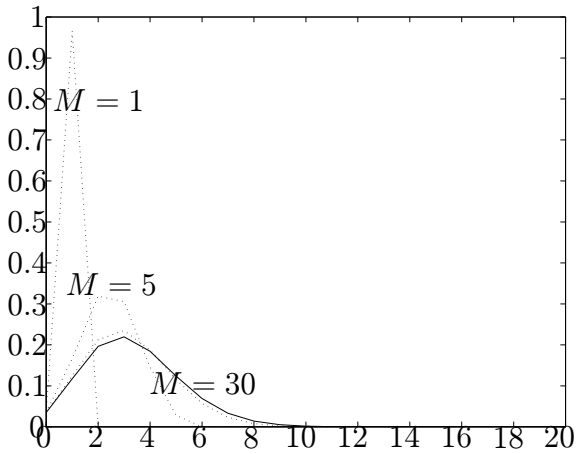
$$\begin{aligned} f_{X^\theta|\theta}(X^\theta_t, t \,|\, \theta_t, \tilde{x}, \tilde{t}) &= \frac{f_X\left(X_t, t \,|\, \tilde{x}, \tilde{t}\right)}{f_\theta\left(\theta_t, t \,|\, \tilde{x}, \tilde{t}\right)} \\ &\approx \frac{1}{\bar{S}}\sum_{s=1}^{S} \phi\left(X^\theta_t, t \,|\, \hat{X}^{\tilde{x}}_{t-h}[s], t-h\right) \; 1_{\theta_t = \hat{\theta}^{\tilde{x}}_{t-h}[s]}. \end{aligned}$$

Variance-reduction techniques can improve the efficiency of the Monte Carlo integration (see Geweke (1996)). Here, antithetic variates are used in simulating the paths of the state vector. That is, with each new pseudo-random Gaussian $\epsilon[s]$ and uniform $u[s]$, the antithetic variates $-\epsilon[s]$ and $1 - u[s]$ are used as a subsequent scenario.

C.2 Simulation of the Target

The highest value that is reached by the intensities λ^U and λ^D in a typical estimated model[29] is 1225. At this value, the next Figure shows that a Bernoulli approximation that allows for only one jump during an FOMC meeting is not accurate. The Bernoulli density for $h = \frac{1}{365}$ overstates the true probability of one jump. If the number of Bernoulli trials during an FOMC meeting is increased so that $h \geq \frac{1}{30}\frac{1}{365}$, the Bernoulli approximation becomes accurate. To economize on the number of simulated steps (and thereby the computation time for the likelihood evaluation), the FOMC meeting day is divided into $M_s + 1$ intervals, where M_s is a number divisable by 5. During 5 subintervals $[t_i, t_{i+1}]$ of length $h = \frac{M_s}{5}\frac{1}{M_s+1}\frac{1}{365}$, jumps are drawn from a Poisson distribution with constant parameter $\lambda^j(t-h)h$ by truncating

[29] The value is taken from λ^U in the unconstrained full model introduced in Section 3.



Approximation of a Poisson density (solid line) with $\lambda = 1225$ for daily data ($t - \tilde{t} = 1/365$) with Bernoulli trials with success probability $p = 1 - \exp(-\lambda h)$, with $h = \frac{1}{M}\frac{1}{365}$, for different choices of M.

the distribution at $\frac{M_s}{5}$ jumps. In the last subinterval of length $h = \frac{1}{M_s+1}$, a Bernoulli discretization is applied. This approximation procedure is equivalent to 31 Bernoulli trials (with appropriately chosen success probability). In the body of the paper, a choice of $M_s = 30$ is called 'subdividing the FOMC meeting into 30 intervals.'

C.3 SML with Macro Variables

The state vector X is now augmented with the macro-related information $M(t) = (m(t), m_F(t))$, and $X^{(\theta,M)}$ denotes the vector consisting of all coordinates of the state X except θ and M. Analogous to the decomposition (C.18), I can write the density of X conditional on the last observation $X_{\tilde{t}} = x$ in the form

$$f_X(X_t, t \,|\, x, \tilde{t}) = f_{\theta,M}(\theta_t, M_t, t \,|\, x, \tilde{t}) f_{X^{(\theta,M)}|\theta,M}(X_t^{(\theta,M)}, t \,|\, \theta_t, M_t, x, \tilde{t}). \tag{C.19}$$

The first term in (C.19) can be written as

$$f_{\theta,M}(\theta_t, M_t, t \,|\, x, \tilde{t}) = f_M(M_t, t \,|\, x, \tilde{t}) f_\theta(\theta_t, t|x, M_t, \tilde{t}).$$

If an FOMC meeting and a macro jump event happen on the same day, the Fed's target decisions are able to condition on the newly released information, as CPI and NPE releases (at 8:30 a.m.) precede FOMC meetings. Since each observation interval $[\tilde{t}, t]$ is chosen so that it does not contain more than one FOMC meeting, the density of M conditional on $X(\tilde{t})$ depends only on $\theta(\tilde{t})$. Moreover, $[\tilde{t}, t]$ is short enough so as to not contain the release of the current month's macro variable together with the analyst survey of forecasts for the next month. For the i-th macroeconomic release, I therefore have

$$f_M(M_t, t \,|\, x, \tilde{t}) = \begin{cases} f_{m_F}(m_F(t), t \,|\, m_F(\tilde{t}), \theta_{\tilde{t}}, \tilde{t}), & \text{if } \tau_F^i \in [\tilde{t}, t], \\ f_m(m(t), t \,|\, m_F(\tilde{t}), \tilde{t}), & \text{if } \tau^i \in [\tilde{t}, t], \\ f_{m_F}(m_F(t) \,|\, m_F(\tilde{t}), \theta_{\tilde{t}}, \tilde{t}) f_m(m(t), t \,|\, m_F(\tilde{t}), \tilde{t}), & \text{if } \tau^i, \tau_F^i \in [\tilde{t}, t]. \end{cases}$$