A Multifactor, Nonlinear, Continuous-Time Model of Interest Rate Volatility

Jacob Boudoukh[a], Christopher Downing[b], Matthew Richardson[c], Richard Stanton[d] and Robert F. Whitelaw[c*]

November 20, 2007

[*a]Arison School of Business, IDC and NBER; [b]Barclays Global Investors; [c]Stern School of Business, NYU and NBER; [d]Haas School of Business, U.C. Berkeley. We would like to thank John Cochrane, Lars Hansen, Chester Spatt, and seminar participants at the New York Federal Reserve, the Federal Reserve Board, Goldman Sachs, Long Term Capital Management, University of North Carolina, U.C. Berkeley, ITAM, the 1997 meetings of the Western Finance Association, the 1998 Utah Winter Finance Conference, and the NBER asset pricing program for helpful comments. Contact: Prof. M. Richardson, NYU, Stern School of Business, 44 W. 4th St., New York, NY 10012, email: mrichar0@stern.nyu.edu, phone: (212) 998-0349.

A Multifactor, Nonlinear, Continuous-Time Model of Interest Rate Volatility

Abstract

This paper presents a general, nonlinear version of existing multifactor models, such as Longstaff and Schwartz (1992). The novel aspect of our approach is that rather than choosing the model parameterization out of "thin air", our processes are *generated* from the data using approximation methods for multifactor continuous-time Markov processes. In applying this technique to the short- and long-end of the term structure for a general two-factor diffusion process for interest rates, a major finding is that the volatility of interest rates is increasing in the level of interest rates *mostly* for sharply upward sloping term structures. In fact, the slope of the term structure plays a larger role in determining the magnitude of the diffusion coefficient. As an application, we show how this method can be used to value fixed-income contingent claims.

1 Introduction

When one sees so many coauthors on a single paper and they are not from the hard sciences, the natural question is why? Looking at both the number and quality of contributors to this volume and how much econometric talent Rob Engle has helped nurture through his career, however, it becomes quite clear that the only way we could participate in Rob's *Festschrift* is to pool our limited abilities in *Financial Econometrics*. Given Rob's obvious importance to econometrics, and in particular to finance via his seminal work on volatility, it is quite humbling to be asked to contribute to this volume.

Looking over Rob's career, it is clear how deeply rooted the finance field is in Rob's work. When one thinks of the major empirical papers in the area of fixed income, Fama and Bliss (1987), Campbell and Shiller (1991), Litterman and Scheinkman (1991), Chan, Karolyi, Longstaff and Sanders (1992), Longstaff and Schwartz (1992), Pearson and Sun (1994), Aït-Sahalia (1996b) and Dai and Singleton (2000) come to mind. Yet in terms of citations, all of these papers are dominated by Rob's 1987 paper with David Lilien and Russell Robins, "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model." In this paper, further expanded upon in Engle and Ng (1993), the authors present evidence that the yield curve is upward sloping when interest rate volatility is high via an ARCH-M effect on term premia. The result is quite natural to anyone who teaches fixed income and tries to relate the tendency for the term structure to be upward sloping to the duration of the underlying bonds. Given this work by Rob, our contribution to this *Festschrift* is to explore the relation between volatility and the term structure more closely.

It is now widely believed that interest rates are affected by multiple factors.[1] Nevertheless, most of our intuition concerning bond and fixed-income derivative pricing comes from stylized facts generated by single-factor, continuous-time interest rate models. For example, the finance literature is uniform in its view that interest rate volatility is increasing in interest rate levels, though there is some disagreement about the rate of increase (see, for example, Chan, Karolyi, Longstaff and Sanders (1992), Aït-Sahalia (1996b), Conley, Hansen, Luttmer and Scheinkman (1995), Brenner, Harjes and Kroner (1996) and Stanton (1997)). If interest rates possess multiple factors such as the level and slope of the term structure (Litterman

[1]See, for example, Stambaugh (1988), Litterman and Scheinkman (1991), Longstaff and Schwartz (1992), Pearson and Sun (1994), Andersen and Lund (1997), Dai and Singleton (2000) and Collin-Dufresne, Goldstein and Jones (2006) to name a few. This ignores the obvious theoretical reasons for multi-factor pricing, as in Brennan and Schwartz (1979), Schaefer and Schwartz (1984), Heath, Jarrow and Morton (1988), Longstaff and Schwartz (1992), Chen and Scott (1992), Duffie and Kan (1996), Ahn, Dittmar and Gallant (2002) and Piazzesi (2005), among others.

and Scheinkman (1991)), and given the Engle, Lilien and Robins (1987) finding, then this volatility result represents an average over all possible term structure slopes. Therefore, conditional on any particular slope, volatility may be severely misestimated, with serious consequences especially for fixed-income derivative pricing.

Two issues arise in trying to generate stylized facts about the underlying continuous-time, stochastic process for interest rates. First, how do we specify *ex ante* the drift and diffusion of the multivariate process for interest rates so that it is consistent with the true process underlying the data? Second, given that we do not have access to continuous-time data, but instead to interest rates/bond prices at discretely sampled intervals, how can we consistently infer an underlying continuous-time multivariate process from these data? In single-factor settings, there has been much headway at addressing these issues (see, for example, Aït-Sahalia (1996a, 2007a), Conley, Hansen, Luttmer and Scheinkman (1995), and Stanton (1997)). Essentially, using variations on nonparametric estimators with carefully chosen moments, the underlying single-factor, continuous-time process can be backed out of interest rate data.

Here, we extend the work of Stanton (1997) to a multivariate setting and provide for the non-parametric estimation of the drift and volatility functions of multivariate stochastic differential equations.[2] Basically, we use Milshtein's (1978) approximation schemes for writing expectations of functions of the sample path of stochastic differential equations in terms of the drift, volatility and correlation coefficients. If the expectations are known (or, in our case, estimated nonparametrically) and the functions are chosen appropriately, then the approximations can be inverted to recover the drift, volatility and correlation coefficients. In this paper, we apply this technique to the short- and long-end of the term structure for a general two-factor, continuous-time diffusion process for interest rates. Our methods can be viewed as a nonparametric alternative to the affine class of multifactor continuous-time interest rate models studied in Longstaff and Schwartz (1992), Duffie and Kan (1996), Dai and Singleton (2000) and Aït-Sahalia and Kimmel (2007b), the quadratic term structure class studied in Ahn, Dittmar and Gallant (2002), and the non-affine parametric specifications of Andersen and Lund (1997). As an application, we show directly how our model relates to the two-factor model of Longstaff and Schwartz (1992).

Our paper provides two contributions to the existing literature. First, in estimating this multi-factor diffusion process, some new empirical facts emerge from the data. Of particular

[2]An exception is Aït-Sahalia (2007b) and Aït-Sahalia and Kimmel (2007b) who provide closed form expansions for the log-likelihood function for a wide class of multivariate diffusions.

note, while the volatility of interest rates increases in the level of interest rates, it does so primarily for sharply upward sloping term structures. Thus, the results of previous studies, suggesting an almost exponential relation between interest rate volatility and levels, are due to the term structure on average being upward sloping, and is not a general result per se. Moreover, our volatility result holds for both the short- and long-term rates of interest. Thus, conditional on particular values of the two factors, such as a high short rate of interest and a negative slope of the term structure, the term structure of interest rate volatilities is generally at a lower level across maturities than implied by previous work.

The second contribution is methodological. In this paper, we provide a way of linking empirical facts and continuous-time modeling techniques so that generating implications for fixed-income pricing is straightforward. Specifically, we use nonparametrically estimated conditional moments of "relevant pricing factors" to build a multifactor continuous-time diffusion process which can be used to price securities. This process can be considered a generalization of the Longstaff and Schwartz (1992) two-factor model. Using this estimated process, we then show how to value fixed-income securities, in conjunction with an estimation procedure for the functional for the market prices of risk. Since the analysis is performed nonparametrically without any priors on the underlying economic structure, the method provides a unique opportunity to study the economic structure's implications for pricing. Of course, ignoring the last twenty-five years of term structure theory and placing more reliance on empirical estimation, with its inevitable estimation error, may not be a viable alternative on its own. Nevertheless, we view this approach as helpful for understanding the relation between interest rate modeling and fixed-income pricing.

2 The Stochastic Behavior of Interest Rates: Some Evidence

In this section, we provide some preliminary evidence for the behavior of interest rates across various points of the yield curve. Under the assumption that there are two interest-rate dependent state variables, and that these variables are spanned by the short rate of interest and the slope of the term structure, we document conditional means and volatilities of changes in the 6-month through 5-year rates of interest. The results are generated nonparametrically, and thus impose no structure on the underlying functional forms for the term structure of interest rates.

2.1 Data Description

Daily values for constant maturity Treasury yields on the 3-year, 5-year and 10-year U.S. government bond were collected from Datastream over the period January 1983 to December 2006. In addition, 3-month, 6-month and 1-year T-bill rates were obtained from the same source, and converted to annualized yields. This provides us with over 6,000 daily observations.

The post-1982 period was chosen because there is considerable evidence that the period prior to 1983 came from a different regime (see, for example, Huizinga and Mishkin (1984), Sanders and Unal (1988), Klemkosky and Pilotte (1992), and Torous and Ball (1995)). In particular, these researchers argue that the October 1979 change in Federal Reserve operating policy led to a once-and-for-all shift in the behavior of the short term riskless rate. Since the Federal Reserve *experiment* ended in November 1982, it is fairly standard to treat only the post late 1982 period as stationary.

In estimating the conditional distribution of the term structure of interest rates, we employ two conditioning factors. These factors are the short rate of interest — defined here as the 3-month yield — and the slope of the term structure — defined as the spread between the 10-year and 3-month yields. These variables are chosen to coincide with interest rate variables used in other studies (see Litterman and Scheinkman (1991) and Chan, Karolyi, Longstaff and Sanders (1992), among others). Figure 1 graphs the time series of both the short rate and spread. Over the 1983 to 2006 period, the short rate ranges from 1% to 11%, while the spread varies from -1% to 4%. There are several distinct periods of low and high interest rates, as well as spread ranges. Since the correlation between the short rate and spread is -0.31, there exists the potential for the two variables combined to possess information in addition to a single factor.

Figure 2 presents a scatter plot of the short rate and term structure slope. Of particular importance to estimating the conditional distribution of interest rates is the availability of the conditioning data. Figure 2 shows that there are two holes in the data ranges, namely at low short rates (i.e., from 1–4%) and low spreads (i.e., from -1–2%), and at high short rates (i.e., from 9.5–11.5%) and low spreads (i.e., from -1–1%). This means that the researcher should be cautious in interpreting the implied distribution of interest rates conditional on these values for the short rate and spread.

2.2 The Conditional Distribution of Interest Rates: A First Look

In order to understand the stochastic properties of interest rates, consider conditioning the data on four possible states: (i) high level (i.e., of the short rate)/high slope, (ii) high level/low slope, (ii) low level/low slope, and (iv) low level/high slope. In a generalized method of moments framework, the moment conditions are:[3]

$$E\begin{pmatrix} (\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{hr:hs}) \times I_{t,hr:hs} \\ (\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{hr:ls}) \times I_{t,hr:ls} \\ (\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{lr:ls}) \times I_{t,lr:ls} \\ (\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{lr:hs}) \times I_{t,lr:hs} \\ [(\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{hr:hs})^2 - {\sigma^{\tau}_{hr:hs}}^2] \times I_{t,hr:hs} \\ [(\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{hr:ls})^2 - {\sigma^{\tau}_{hr:ls}}^2] \times I_{t,hr:ls} \\ [(\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{lr:ls})^2 - {\sigma^{\tau}_{lr:ls}}^2] \times I_{t,lr:ls} \\ [(\Delta i^{\tau}_{t,t+1} - \mu^{\tau}_{lr:hs})^2 - {\sigma^{\tau}_{lr:hs}}^2] \times I_{t,lr:hs} \end{pmatrix} = 0, \tag{1}$$

where $\Delta i^{\tau}_{t,t+1}$ is the change in the τ-period interest rate from t to $t+1$, $\mu^{\tau}_{\cdot|\cdot}$ is the mean change in rates conditional on one of the four states occurring, $\sigma^{\tau}_{\cdot|\cdot}$ is the volatility of the change in rates conditional on these states, and $I_{t,\cdot|\cdot} = 1$ if $[\cdot|\cdot]$ occurs, zero otherwise. These moments, μ^{τ} and σ^{τ}, thus represent coarse estimates of the underlying conditional moments of the distribution of interest rates.

These moment conditions allow us to test a variety of restrictions. First, are $\sigma^{\tau}_{hr:hs} = \sigma^{\tau}_{hr:ls}$ and $\sigma^{\tau}_{lr:hs} = \sigma^{\tau}_{lr:ls}$? That is, does the slope of the term structure help explain volatility at various interest rate levels? Second, similarly, with respect to the mean, are $\mu^{\tau}_{hr:hs} = \mu^{\tau}_{hr:ls}$ and $\mu^{\tau}_{lr:hs} = \mu^{\tau}_{lr:ls}$? Table 1 provides estimates of $\mu^{\tau}_{\cdot|\cdot}$ and $\sigma^{\tau}_{\cdot|\cdot}$, and the corresponding test statistics. Note that the framework allows for autocorrelation and heteroskedasticity in the underlying squared interest rate series when calculating the variance-covariance matrix of the estimates. Further, the cross-correlation between the volatility estimates is taken into account in deriving the test statistics.

Several facts emerge from Table 1. First, as documented by others (e.g., Chan, Karolyi, Longstaff and Sanders (1992), and Aït-Sahalia (1996a)), interest rate volatility is increasing in the short rate of interest. Of some interest here, this result holds across the yield curve. That is, conditional on either a low or high slope, volatility is higher for the 6-month, 1-year, 3-year and 5-year rates at higher levels of the short rate. Second, the slope also plays

[3] We define a low (high) level or spread as one that lies below (above) its unconditional mean. Here, this mean is being treated as a known constant, though, of course, it is estimated via the data.

an important role in determining interest rate volatility. In particular, at high levels of interest rates, the volatility of interest rates across maturities is much higher at steeper slopes. For example, the 6-month and 5-year volatilities rise from 5.25 and 6.35 to 7.65 and 7.75 basis points, respectively. Formal tests of the hypothesis $\sigma^{\tau}_{hr:hs} = \sigma^{\tau}_{hr:ls}$ provide 1% level rejections at each of the maturities. There is some evidence in the literature that expected returns on bonds are higher for steeper term structures (see, for example, Fama (1986) and Boudoukh, Richardson, Smith and Whitelaw (1999a, 1999b)); these papers and the finding of Engle, Lilien and Robins (1987) may provide a link to the volatility result here. Third, the effect of the slope is most important at high interest rate levels. At low short rate levels, though the volatility at low slopes is less than that at high slopes, the effect is much less pronounced. This is confirmed by the fact that a number of the p-values are no longer significant at conventional levels for the test of the hypothesis, $\sigma^{\tau}_{lr:hs} = \sigma^{\tau}_{lr:ls}$. Fourth, the conditional means, though not in general reliably estimated, are consistent with existing results in the literature (e.g., Chan, Karolyi, Longstaff and Sanders (1992), Aït-Sahalia (1996a), and Stanton (1997)). That is, at low levels of interest rates, the mean tends to be greater than at high interest rates, which can be explained by mean reversion. However, the table also provides an interesting new result, namely that the effect of the slope is of higher magnitude than the level. Further, low slopes tend to be associated with negative changes in rates, while high slopes are linked to positive interest rate changes.

2.3 The Conditional Distribution of Interest Rates: A Closer Look

In order to generalize the results of Section 2.2, we employ a kernel estimation procedure for estimating the relation between interest rate changes and components of the term-structure of interest rates. Kernel estimation is a nonparametric method for estimating the joint density of a set of random variables. Specifically, given a time series $\Delta i^{\tau}_{t,t+1}$, i^r_t and i^s_t (where i^r is the level of interest rates, and i^s is the slope), generated from an unknown density $f(\Delta i^{\tau}, i^r, i^s)$, then a kernel estimator of this density is

$$\hat{f}(\Delta i^{\tau}, i^r, i^s) = \frac{1}{Th^m} \sum_{t=1}^{T} K\left(\frac{(\Delta i^{\tau}, i^r, i^s) - (\Delta i^{\tau}_{t,t+1}, i^r_t, i^s_t)}{h} \right), \tag{2}$$

where $K(\cdot)$ is a suitable kernel function and h is the window width or smoothing parameter.

We employ the commonly used independent multivariate normal kernel for $K(\cdot)$. The other parameter, the window width, is chosen based on the dispersion of the observations.

For the independent multivariate normal kernel, Scott (1992) suggests the window width,

$$\hat{h} = k\hat{\sigma}_i T^{\frac{-1}{m+4}},$$

where $\hat{\sigma}_i$ is the standard deviation estimate of each variable z_i, T is the number of observations, m is the dimension of the variables, and k is a scaling constant often chosen via cross-validation. Here, we employ a cross-validation procedure to find the k which provides the right trade-off between the bias and variance of the errors. Across all the data points, we find the k's which minimize the mean-squared error between the observed data and the estimated conditional data. This mean-squared error minimization is implemented using a Jackknife-based procedure. In particular, the various implied conditional moments at each data point are estimated using the entire sample, except for the actual data point and its nearest neighbors.[4] Once the k is chosen, the actual estimation of the conditional distribution of interest rates involves the entire sample, albeit using window widths chosen from partial samples. To coincide with Section 2.2., we focus on the first two conditional moments of the distribution, and it is possible to show that

$$\hat{\mu}_{\Delta i^\tau}(i^r, i^s) = \sum_{t=1}^{T} w_t(i^r, i^s)\Delta i_t^\tau \tag{3}$$

$$\hat{\sigma^2}_{\Delta i^\tau}(i^r, i^s) = \sum_{t=1}^{T} w_t(i^r, i^s)(\Delta i_t^\tau - \hat{\mu}_{\Delta i^\tau}(i^r, i^s))^2, \tag{4}$$

where $w_t(i^r, i^s) = K\left(\frac{(i^r,i^s)-(i_t^r,i_t^s)}{h}\right) / \sum_{t=1}^{T} K\left(\frac{(i^r,i^s)-(i_t^r,i_t^s)}{h}\right)$. The weights, $w_t(i^r, i^s)$, are determined by how close the chosen state, i.e., the particular values of the level and slope, i^r and i^s, is to the observed level and slope of the term structure, i_t^r and i_t^s.

As an illustration, using equation (4), Figure 3 provide estimates of the volatility of daily changes in the 1-year rate, conditional on the current level of the short rate and the slope of the term structure (i.e., i_t^r and i_t^s). While Figure 3 represents only the 1-year rate, the same effects carry through to the rest of the yield curve and have therefore been omitted for purposes of space. The figure maps these estimates to the relevant range of the data, in particular, for short rates ranging from 3% to 11% and slopes ranging from 0.0% to 3.5%. The volatility is maximized at high interest rate levels and high slopes though the more

[4]Due to the serial dependence of the data, we performed the cross-validation omitting 100 observations, i.e., four months in either direction of the particular data point in question. Depending on the moments in question, the optimal k's range from roughly 1.7 to 27.6, which implies approximately twice to twenty eight times the smoothing parameter of Scott's asymptotically optimal implied value.

dramatic changes occur at high slopes.

To see this a little more clearly, Figures 4 and 5 present cut-throughs of Figure 3 across the term structure at short rates of 8.0% and 5.5%, respectively. From Figure 2, these levels represent data ranges in which there are many different slopes; thus, conditional on these levels, the estimated relation between the volatility of the 6-month, 1-year, 3-year and 5-year rates as a function of the slope is more reliable. Several observations are in order. First, as seen from the figures, volatility is increasing in the slope for all maturities, though primarily only for steep term structures, i.e., above 2.0%. Second, volatility is also higher at greater magnitudes of the short rate albeit less noticeably. These results suggest that any valuation requiring a volatility estimate of interest rates should be done with caution. For example, estimating volatility when the term structure is flat relative to upward sloping should lead to quite different point estimates. Third, the relation between volatility and the slope is nonlinear, which, as it turns out in Section 3.3, will lead to a nonlinear continuous-time diffusion process. This feature can be potentially important as the majority of the multifactor, term structure pricing models are derived from the affine class. Alternatively, Figures 6 and 7 provide cut-throughs of Figure 3 across the term structure at slopes of 2.75% and 1.00%, respectively. These slopes represent data ranges in which there are a number of observations of the interest rate level. The figures show that the estimated relation between the volatility of the 6-month, 1-year, and especially the 3-year and 5-year rates as a function of the level depends considerably on the slope of the term structure. For example, the volatility of the 6-mth and 1-year interest rate change are almost flat over levels of 3.0% to 6.0% at low slopes, whereas it increases roughly 200 basis points at high slopes. Similarly, even at the long end of the yield curve, the increase in volatility is higher at high versus low slopes.

3 Estimation of a Continuous-Time Multi-factor Diffusion Process

The results of Section 2 suggest that the volatility of changes in the term structure of interest rates depends on at least two factors. Given the importance of continuous-time mathematics in the fixed income area, the question arises as to how these results can be interpreted in a continuous-time setting. Using data on bond prices, and explicit theoretical pricing models (e.g., Cox, Ingersoll and Ross (1985)), Brown and Dybvig (1989), Pearson and Sun (1994),

Gibbons and Ramaswamy (1994) and Dai and Singleton (2000) all estimate parameters of the underlying interest-rate process in a fashion consistent with the underlying continuous-time model. These procedures limit themselves, however, to fairly simple specifications.

As a result, a literature emerged which allows estimation and inference of fairly general continuous-time diffusion processes using discretely sampled data. Aït-Sahalia (2007a) provides a survey of this literature and we provide a quick review here. First, at a parametric level, there has been considerable effort in the finance literature at working through maximum likelihood applications of continuous-time processes with discretely sampled data, starting with Lo (1988) and continuing more recently with Aït-Sahalia (2002) and Aït-Sahalia and Kimmel (2007a, 2007b). Second, by employing the infinitesimal generators of the underlying continuous-time diffusion processes, Hansen and Scheinkman (1995) and Conley, Hansen, Luttmer and Scheinkman (1995) construct moment conditions that also make the investigation of continuous-time models possible with discrete time data. Third, in a nonparametric framework, Aït-Sahalia (1996a, 1996b) develops a procedure for estimating the underlying process for interest rates using discrete data by choosing a model for the drift of interest rates and then nonparametrically estimating its diffusion function. Finally, as an alternative method, Stanton (1997) employs approximations to the true drift and diffusion of the underlying process, and then nonparametrically estimates these approximation terms to back out the continuous-time process (see also Bandi (2002), Chapman and Pearson (2000) and Pritsker (1998)). The advantage of this approach is twofold: (i) similar to the other procedures, the data need only be observed at discrete time intervals, and (ii) the drift and diffusion are unspecified, and thus may be highly nonlinear in the state variable.

In this section, we extend the work of Stanton (1997) to a multivariate setting and provide for the non-parametric estimation of the drift and volatility functions of multivariate stochastic differential equations. Similar to Stanton (1997), we use Milshtein's (1978) approximation schemes for writing expectations of functions of the sample path of stochastic differential equations in terms of the drift and volatility coefficients. If the expectations are known (albeit estimated nonparametrically in this paper) and the functions are chosen appropriately, then the approximations can be inverted to recover the drift and volatility coefficients.

3.1 Drift, Diffusion and Correlation Approximations

Assume that no arbitrage opportunities exist, and that bond prices are functions of two state variables, the values of which can always be inverted from the current level, R_t, and a second state variable, S_t. Assume that these variables follow the (jointly) Markov diffusion process

$$dR_t = \mu_R(R_t, S_t)\,dt + \sigma_R(R_t, S_t)\,dZ_t^R \tag{5}$$

$$dS_t = \mu_S(R_t, S_t)\,dt + \sigma_S(R_t, S_t)\,dZ_t^S, \tag{6}$$

where the drift, volatility and correlation coefficients (i.e., the correlation between Z^R and Z^S) all depend on R_t and S_t. Define the vector $X_t = (R_t, S_t)$.

Under suitable restrictions on μ, σ, and a function f, we can write the conditional expectation $E_t\left[f(X_{t+\Delta})\right]$ in the form of a Taylor series expansion,[5]

$$E_t\left[f(X_{t+\Delta})\right] = f(X_t) + \mathcal{L}f(X_t)\Delta + \frac{1}{2}\mathcal{L}^2 f(X_t)\Delta^2 + \ldots + \frac{1}{n!}\mathcal{L}^n f(X_t)\Delta^n + O(\Delta^{n+1}), \tag{7}$$

where $\mathcal{L}$ is the infinitesimal generator of the multivariate process $\{X_t\}$ (see Øksendal (1985) and Hansen and Scheinkman (1995)), defined by

$$\mathcal{L}f(X_t) = \left(\frac{\partial f(X_t)}{\partial X_t}\right)\mu_X(X_t) + \frac{1}{2}\text{trace}\left[\Sigma(X_t)\left(\frac{\partial^2 f(X_t)}{\partial X_t \partial X_t'}\right)\right],$$

where

$$\Sigma(X_t) = \begin{pmatrix} \sigma_R^2(R_t, S_t) & \rho(R_t, S_t)\sigma_R(R_t, S_t)\sigma_S(R_t, S_t) \\ \rho(R_t, S_t)\sigma_R(R_t, S_t)\sigma_S(R_t, S_t) & \sigma_S^2(R_t, S_t) \end{pmatrix}.$$

Equation (7) can be used to construct numerical approximations to $E_t[f(X_{t+\Delta})]$ in the form of a Taylor series expansion, given known functions μ_R, μ_S, ρ, σ_R and σ_S (see, for

[5]For a discussion see, for example, Hille and Phillips (1957), Chapter [11]. Milshtein (1974, 1978) gives examples of conditions under which this expansion is valid, involving boundedness of the functions μ, σ, f and their derivatives. There are some stationary processes for which this expansion does *not* hold for the functions f that we shall be considering, including processes such as

$$dx = \mu\,dt + x^3\,dZ,$$

which exhibit "volatility induced stationary" (see Conley, Hansen, Luttmer and Scheinkman (1995)). However, any process for which the first order Taylor series expansion fails to hold (for linear f) will also fail if we try to use the usual numerical simulation methods (e.g. Euler discretization). This severely limits their usefulness in practice.

example, Milshtein (1978)). Alternatively, given an appropriately chosen set of functions $f(\cdot)$ and nonparametric estimates of $E_t[f(X_{t+\Delta})]$, we can use equation (7) to construct approximations to the drift, volatility and correlation coefficients (i.e., μ_R, μ_S, ρ, σ_R and σ_S) of the underlying multifactor, continuous-time diffusion process. The nice feature of this method is that the functional forms for μ_R, μ_S, ρ, σ_R and σ_S are quite general, and can be estimated nonparametrically from the underlying data. Rearranging equation (7), and using a time step of length $i\Delta$ ($i = 1, 2, \ldots$), we obtain

$$\begin{aligned} \widehat{E}^i(X_t) &\equiv \frac{1}{i\Delta} E_t\left[f(X_{t+i\Delta}) - f(X_t)\right], \\ &= \mathcal{L}f(X_t) + \frac{1}{2}\mathcal{L}^2 f(X_t)(i\Delta) + \ldots + \frac{1}{n!}\mathcal{L}^n f(X_t)(i\Delta)^{n-1} + O(\Delta^n). \end{aligned} \tag{8}$$

From equation (8), each of the $\widehat{E}^i$ is a first order approximation to $\mathcal{L}f$,

$$\widehat{E}^i(X_t) = \mathcal{L}f(X_t) + O(\Delta).$$

Now consider forming linear combinations of these approximations, $\sum_{i=1}^N \alpha_i \widehat{E}^i(X_t)$. That is, from equation (8),

$$\sum_{i=1}^N \alpha_i \widehat{E}^i(X_t) = \left[\sum_{i=1}^N \alpha_i\right] \mathcal{L}f(X_t) + \frac{1}{2}\left[\sum_{i=1}^N \alpha_i i\right] \mathcal{L}^2 f(X_t)\Delta + \frac{1}{6}\left[\sum_{i=1}^N \alpha_i i^2\right] \mathcal{L}^3 f(X_t)\Delta^2 + \ldots. \tag{9}$$

Can we choose the α_i so that this linear combination is an approximation to $\mathcal{L}f$ of order N?

For the combination to be an approximation to $\mathcal{L}f$, we require first that the weights $\alpha_1, \alpha_2, \ldots, \alpha_N$ sum to 1. Furthermore, from equation (9), in order to eliminate the first order error term, the weights must satisfy the equation

$$\sum_{i=1}^N \alpha_i i = 0.$$

More generally, in order to eliminate the nth order error term ($n \leq N-1$), the weights must satisfy the equation,

$$\sum_{i=1}^N \alpha_i i^n = 0.$$

We can write this set of restrictions more compactly in matrix form as

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 3 & \cdots & N \\ 1 & 4 & 9 & \cdots & N^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2^{N-1} & 3^{N-1} & \cdots & N^{N-1} \end{pmatrix} \alpha \equiv V\alpha = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The matrix V is called a Vandermonde matrix, and is invertible for any value of N. We can thus obtain α by calculating

$$\alpha = V^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0. \end{pmatrix}. \tag{10}$$

For example, for $N = 3$, we obtain

$$\alpha = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{11}$$

$$= \begin{pmatrix} 3 \\ -3 \\ 1 \end{pmatrix}. \tag{12}$$

Substituting α into equation (9), and using equation (8), we get the following third order approximation of the infinitesimal generator of the process $\{X_t\}$:

$$\mathcal{L}f(X_t) = \frac{1}{6\Delta}\left[18E_t\left(f(X_{t+\Delta}) - f(X_t)\right) - 9E_t\left(f(X_{t+2\Delta}) - f(X_t)\right) + 2E_t\left(f(X_{t+3\Delta}) - f(X_t)\right)\right] + O(\Delta^3).$$

To approximate a particular function $g(x)$, we now need merely to find a specific function f satisfying

$$\mathcal{L}f(x) = g(x).$$

For our purposes, consider the functions

$$\begin{aligned} f_{(1)}(R) &\equiv R - R_t, \\ f_{(2)}(S) &\equiv S - S_t, \\ f_{(3)}(R) &\equiv (R - R_t)^2, \\ f_{(4)}(S) &\equiv (S - S_t)^2, \\ f_{(5)}(R,S) &\equiv (R - R_t)(S - S_t). \end{aligned}$$

From the definition of $\mathcal{L}$, we have

$$\begin{aligned} \mathcal{L}f_{(1)}(R) &= \mu_R(R,S), \\ \mathcal{L}f_{(2)}(S) &= \mu_S(R,S), \\ \mathcal{L}f_{(3)}(R) &= 2(R - R_t)\mu_R(R,S) + \sigma_R^2(R,S), \\ \mathcal{L}f_{(4)}(S) &= 2(S - S_t)\mu_S(R,S) + \sigma_S^2(R,S), \\ \mathcal{L}f_{(5)}(R,S) &= (S - S_t)\mu_R(R,S) + (R - R_t)\mu_S(R,S) + \rho(R,S)\sigma_R(R,S)\sigma_S(R,S). \end{aligned}$$

Evaluating these at $R = R_t$, $S = S_t$, we obtain

$$\begin{aligned} \mathcal{L}f_{(1)}(R_t) &= \mu_R(R_t,S_t), \\ \mathcal{L}f_{(2)}(S_t) &= \mu_S(R_t,S_t), \\ \mathcal{L}f_{(3)}(R_t) &= \sigma_R^2(R_t,S_t), \\ \mathcal{L}f_{(4)}(S_t) &= \sigma_S^2(R_t,S_t), \\ \mathcal{L}f_{(5)}(R_t,S_t) &= \rho(R_t,S_t)\sigma_R(R_t,S_t)\sigma_S(R_t,S_t). \end{aligned}$$

Using each of these functions in turn as the function f above, we can generate approximations to μ_R, μ_S, σ_R, σ_S and ρ respectively. For example, the third order approximations (taking square roots for σ_R and σ_S) are

$$\begin{aligned} \mu_R(R_t,S_t) &= \frac{1}{6\Delta}\left[18E_t\left(R_{t+\Delta} - R_t\right) - 9E_t\left(R_{t+2\Delta} - R_t\right) + 2E_t\left(R_{t+3\Delta} - R_t\right)\right] \\ &\qquad + O(\Delta^3), \qquad (13) \\ \mu_S(R_t,S_t) &= \frac{1}{6\Delta}\left[18E_t\left(S_{t+\Delta} - S_t\right) - 9E_t\left(S_{t+2\Delta} - S_t\right) + 2E_t\left(S_{t+3\Delta} - S_t\right)\right] \\ &\qquad + O(\Delta^3), \\ \sigma_R(R_t,S_t) &= \sqrt{\frac{1}{6\Delta}\left(18E_t\left[\left(R_{t+\Delta} - R_t\right)^2\right] - 9E_t\left[\left(R_{t+2\Delta} - R_t\right)^2\right] + 2E_t\left[\left(R_{t+3\Delta} - R_t\right)^2\right]\right)} \end{aligned}$$

$$\begin{aligned}
\sigma_S(R_t, S_t) &= \sqrt{\frac{1}{6\Delta}\left(18E_t\left[(S_{t+\Delta}-S_t)^2\right]-9E_t\left[(S_{t+2\Delta}-S_t)^2\right]+2E_t\left[(S_{t+3\Delta}-S_t)^2\right]\right)} \\
\sigma_{RS}(R_t, S_t) &= \frac{1}{6\Delta}\left(18E_t\left[(R_{t+\Delta}-R_t)(S_{t+\Delta}-S_t)\right]-9E_t\left[(R_{t+2\Delta}-R_t)(S_{t+2\Delta}-S_t)\right]\right. \\
&\qquad \left.+2E_t\left[(R_{t+3\Delta}-R_t)(S_{t+3\Delta}-S_t)\right]\right).
\end{aligned}$$

The approximations of the drift, volatility and correlation coefficients are written in terms of the true first, second and cross moments of multiperiod changes in the two state variables. If the two-factor assumption is appropriate, and a large stationary time series is available, then these conditional moments can be estimated using appropriate nonparametric methods. In this paper, we estimate the moments using multivariate density estimation, with appropriately chosen factors as the conditioning variables. All that is required is that these factors span the same space as the true state variables.[6] The results for daily changes were provided in Section 2. Equation (13) shows that these estimates are an important part of the approximations to the underlying continuous-time dynamics. By adding multiperiod extensions of these nonparametric estimated conditional moments, we can estimate the drift, volatility and correlation coefficients of the multifactor process described by equations (5) and (6).

Figure 8 provides the first, second and third order approximations to the diffusion of the short rate against the short rate level and the slope of the term structure.[7] The most notable result is that a first order approximation works well; thus, one can consider the theoretical results of this section as a justification for discretization methods currently used in the literature. The description of interest rate behavior given in Section 2, therefore, carries through to the continuous-time setting. Our major finding is that the volatility of interest rates is increasing in the level of interest rates *mostly* for sharply upward sloping term structures. The question then is what does Figure 8, and more generally the rest of the estimated process, mean for fixed-income pricing?

[6] See Duffie and Kan (1996) for a discussion of the conditions under which this is possible (in a linear setting).

[7] Figures showing the various approximations to the drift of the short rate, the drift and diffusion of the slope, and the correlation between the short rate and the slope are available upon request.

4 A Generalized Longstaff and Schwartz (1992) Model

Longstaff and Schwartz (1992) provide a two-factor general equilibrium model of the term structure. Their model is one of the more popular versions within the affine class of models for describing the yield curve (see also Cox, Ingersoll and Ross (1985), Chen and Scott (1995), Duffie and Kan (1996) and Dai and Singleton (2000)). In the Longstaff and Schwartz setting, all fixed-income instruments are functions of two fundamental factors, the instantaneous interest rate and its volatility. These factors follow diffusion processes, which in turn lead to a fundamental valuation condition for the price of any bond, or bond derivative. As an alternative, here we also present a two-factor continuous-time model for interest rates. The results of Section 2 suggest that the affine class may be too restrictive.

While our results shed valuable light on the factors driving interest rate movements, however, there are potential problems in using this specification to price interest rate contingent claims. A general specification for R_t and S_t (and the associated prices of risk) may allow arbitrage opportunities if either of these state variables is a known function of an asset price.[8] Of course, this point is true of all previous estimations of continuous-time processes to the extent that they use a priced proxy as the instantaneous rate. If we are willing to assume that we have the right factors, however, then there is no problem in an asymptotic sense. That is, since we are estimating these processes nonparametrically, as the sample size gets larger, our estimates will converge to the true functions, which are automatically arbitrage-free (if the economy is). Nevertheless, this is of little consolation if we are trying to use the *estimated* functions to price assets.

To get around this problem, we need to write the model in a form in which neither state variable is an asset price or a function of asset prices. In this paper, we follow convention by using the observable 3-month yield as a proxy for the instantaneous rate, R_t. Furthermore, suppose that the mapping from (R, S) to (R, σ_R) is invertible,[9] so we can write asset prices as a function of R and σ_R, instead of R and S.[10] Since σ_R is not an asset price, using this

[8] See, for example, Duffie, Ma and Yong (1995). The problem is that, given such a model, we can price any bond, and are thus able to calculate what the state variable "ought" to be. Without imposing any restrictions on the assumed dynamics for R_t and S_t, there is no guarantee that we will get back to the same value of the state variable that we started with.

[9] That is, for a given value of R_t, the volatility, σ_R, is monotonic in the spread, S. This is the case in most existing multifactor interest rate models, including, for example all affine models, such as Longstaff and Schwartz (1992).

[10] This follows by writing

$$V(R, S) = V(R, S(R, \sigma_R)) \equiv U(R, \sigma_R).$$

variable avoids the inconsistency problem.

Specifically, suppose that the true model governing interest rate movements is a generalization of the two factor Longstaff and Schwartz (1992) model,

$$dR_t = \mu_R(R,\sigma)\,dt + \sigma\,dZ_1, \tag{14}$$

$$d\sigma_t = \mu_\sigma(R,\sigma)dt + \rho(R,\sigma)s(R,\sigma)\,dZ_1 + \sqrt{1-\rho^2}s\,dZ_2, \tag{15}$$

where $dZ_1\,dZ_2 = 0$.[11] In vector terms,

$$d(R_t,\sigma_t) = M\,dt + \theta\,dZ,$$

where

$$M \equiv \begin{pmatrix} \mu_R \\ \mu_\sigma \end{pmatrix},$$

$$\theta \equiv \begin{pmatrix} \sigma & 0 \\ \rho s & \sqrt{1-\rho^2}s \end{pmatrix}.$$

Asset prices, and hence the slope of the term structure, can be written as some function of the short rate and instantaneous short rate volatility, $S(R,\sigma)$.

From equations (14) and (15), how do we estimate the underlying processes for R and σ given the estimation results of Section 3? Although the short rate volatility, σ, is not directly observable, it is possible to estimate this process. Specifically, using Ito's Lemma, together with estimates for μ_R, σ_R, μ_S, σ_S and ρ, it is possible to write

$$d\sigma_t = \sigma_R dR_t + \sigma_S dS_t + \frac{1}{2}\left[\sigma_{RR}\sigma^2(R_t,S_t) + \sigma_{SS}\sigma_S^2(R_t,S_t) + 2\sigma_{RS}\sigma(R_t,S_t)\sigma_S(R_t,S_t)\rho(R_t,S_t)\right]dt.$$

Given this equation, and the assumption that the function $S(R,\sigma)$ is invertible, the dynamics of σ_t can be written as a function of the current level of R and σ in a straightforward way.

This procedure requires estimation of a matrix of second derivatives. Although there are well-known problems in estimating higher-order derivatives using kernel density estimation techniques, it is possible to link the results of Section 2 and 3 to this generalized Longstaff and Schwartz (1992) model. In particular, using estimates of the second derivatives (not shown),

[11] This specification is the most convenient to deal with, since we now have orthogonal noise terms. The correlation between the diffusion terms is ρ, and the overall variance of σ is $s^2\,dt$.

several facts emerge. First, due to the small magnitudes of the estimated drifts of the state variables R and S, the drift of σ depends primarily on the second order terms. Consequently, the importance of the second factor (the slope) is determined by how much the sensitivity of short rate volatility to this factor changes relative to the changes in the sensitivity to the first factor (the level). The general pattern is that volatility increases at a slower rate for high levels and a faster rate for high slopes. Consequently, for high volatilities and levels, the drift of volatility is negative, generating mean reversion. The effect of the second factor however is to counter this phenomenon. Second, the diffusion of σ is determined by the sensitivities of short rate volatility to the two factors and the magnitudes of the volatilities of the factors. Based on the estimates of the volatilities and derivatives, the slope has the dominant influence on this effect. In particular, the volatility of σ is high for upward sloping term structures, which also correspond to states with high short rate volatility. Moreover, sensitivity of this diffusion to the two factors is larger in the slope direction than in the level direction.

As an alternative to the above method, we can estimate an implied series for σ by assuming that the function $S(R, \sigma)$ is invertible, i.e., that we can equivalently write the model in the form

$$dR_t = \mu_R(R_t, S_t)dt + \sigma(R_t, S_t)dZ_1^*$$
$$dS_t = \mu_S(R_t, S_t)dt + \sigma_S(R_t, S_t)dZ_2^*,$$

where Z_1^* and Z_2^* may be correlated. To estimate the function $\sigma(R, S)$, we apply the methodology described in Section 3.1 to the function $f_{(3)}(R, S) \equiv (R - R_t)^2$. Applying the estimated function to each observed (R, S) pair in turn yields a series for the volatility σ, which we can then use in estimating the generalized Longstaff and Schwartz (1992) model given in equations (14) and (15).[12] This procedure is in stark contrast to that of Longstaff and Schwartz (1992), and others, who approximate the dynamics of the volatility factor as a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) process. The GARCH process is not strictly compatible with the underlying dynamics of their continuous-time model; here, the estimation is based on approximation schemes to the diffusion process and is internally consistent. Due to the difficulties in estimating derivatives, we choose this second approach

[12] Although the use of an estimated series for σ rather than the true series may not be the most efficient approach, this procedure is consistent. That is, the problem will disappear as the sample size becomes large, and our pointwise estimates of σ converge to the true values.

to estimate the continuous-time process.[13]

4.1 A General Two Factor Diffusion Process: Empirical Results

Figures 10–12 show approximations to equation (15) for the generalized Longstaff and Schwartz (1992) process as a function of the two factors, the instantaneous short rate and its volatility. It is important to point out that there is little available data at low short rates/high volatilities and high short rates/low volatilities, which corresponds to the earlier comment about interest rates and spreads (see Figure 9). Therefore, results in these regions need to be treated cautiously.

Figures 10 and 11 provide the estimates of the continuous-time process for the second interest factor, namely its volatility. Several observations are in order. First, there is estimated mean-reversion in volatility; at low (high) levels of volatility, volatility tends to drift upward (downward). The effect of the level of interest rates on this relation appears minimal. Second, and perhaps most important, there is clear evidence that the diffusion of the volatility process is increasing in the level of volatility, yet is affected by the level of interest rates only marginally. Moreover, volatility's effect is nonlinear in that it takes effect only at higher levels. This finding suggests extreme caution should be applied when inputting interest rate volatility into derivative pricing models. Most of our models take the relation between the level and volatility for granted; however, with increases from 3% to 11% in the interest rate level, both volatility's drift and diffusion exhibit only mild increases. On the other hand, changes in the volatility level of much smaller magnitudes have a much larger impact on the volatility process. This finding links the term structure slope result documented earlier in the paper to a second factor, namely the volatility of the instantaneous rate, and provides a close connection to the Engle, Lilien and Robins (1987) paper mentioned throughout this article.

As the final piece of the multifactor process for interest rates, Figure 12 graphs a first order approximation of the correlation coefficient between the short rate and the volatility, given values of the two factors. Taken at face value, the results suggest a complex variance-covariance matrix between these series in continuous-time. In particular, while the correlation decreases in the volatility for most interest rate levels, there appears to be some nonmonotonicity across the level itself. Why is correlation falling as volatility increases?

[13]Though the first approach provides similar results to the second approach, the functional forms underlying the second method are more smooth and thus more suitable for analysis.

Perhaps, high volatility, just like the corresponding high term structure slope, is associated with aggregate economic phenomena that are less related to the level of interest rates. Given that interest rates are driven by two relatively independent economic factors, namely expectations about both real rates and inflation, this argument seems reasonable. It remains an open question, however, what the exact relation is between Figure 12 and these economic factors.

4.2 Valuation of Fixed-Income Contingent Claims

Given the interest rate model described in equation (15), we can write the price of an interest rate contingent claim as $V(r, \sigma, t)$, depending only on the current values of the two state variables plus time. Then, by Ito's Lemma,

$$\frac{dV(r,\sigma,t)}{V(r,\sigma,t)} = m(r,\sigma,t)\,dt + s_1(r,\sigma,t)\,dZ_1 + s_2(r,\sigma,t)\,dZ_2, \tag{16}$$

where

$$\begin{aligned} m(r,\sigma,t)\,V &= V_t + \mu_r(r,\sigma)V_R + \mu_\sigma(r,\sigma)V_\sigma + \frac{1}{2}\text{trace}\left[\theta^T\,\nabla^2 V(r,\sigma)\,\theta\right], \\ &= V_t + \mu_r(r,\sigma)V_r + \mu_\sigma(r,\sigma)V_\sigma + \frac{1}{2}\sigma^2 V_{rr} + \frac{1}{2}s^2 V_{\sigma\sigma} + \rho\sigma s V_{r\sigma}, \qquad (17) \\ s_1(r,\sigma,t)\,V &= \sigma V_r + \rho s V_\sigma, \\ s_2(r,\sigma,t)\,V &= \sqrt{1-\rho^2}\,s V_\sigma. \end{aligned}$$

The volatility of the asset, σ_V, is given by

$$\begin{aligned} \sigma_V\,V &= \sqrt{\left(\sigma V_r + \rho s V_\sigma\right)^2 + \left(1-\rho^2\right)s^2 V_\sigma^2}, \\ &= \sqrt{\sigma^2 V_r^2 + 2\rho\sigma s V_r V_\sigma + s^2 V_\sigma^2}. \end{aligned}$$

With a one factor interest rate model, to prevent arbitrage, the risk premium on any asset must be proportional to its standard deviation.[14] Similarly, with two factors, absence of arbitrage requires the excess return on an asset to be a linear combination of its exposure

[14]Suppose this did not hold for two risky assets. We could then create a riskless portfolio of these two assets with a return strictly greater than r, leading to an arbitrage opportunity (see Ingersoll (1987)).

to the *two* sources of risk. Thus, if the asset pays out dividends at rate d, we can write

$$m = r - \frac{d}{V} + \lambda_r(r,\sigma)\frac{V_r}{V} + \lambda_\sigma(r,\sigma)\frac{V_\sigma}{V}, \tag{18}$$

where λ_r and λ_σ are the prices of short rate risk and volatility risk respectively. Substituting equation (18) into equation (17), and simplifying, leads to a partial differential equation that must be satisfied by any interest rate contingent claim, assuming the usual technical smoothness and integrability conditions (see, for example, Duffie (1988)),

$$\frac{1}{2}\sigma^2 V_{rr} + [\mu_r - \lambda_r]\, V_r + \frac{1}{2}s^2 V_{\sigma\sigma} + [\mu_\sigma - \lambda_\sigma]\, V_\sigma + \rho\sigma s V_{r\sigma} + V_t - rV + d = 0, \tag{19}$$

subject to appropriate boundary conditions. To price interest rate dependent assets, we need to know not only the processes governing movements in r and σ, but also the prices of risk, λ_r and λ_σ.

Equation (18) gives an expression for these functions in terms of the partial derivatives V_r and V_σ, which could be used to estimate the prices of risk, given estimates of these derivatives for two different assets, plus estimates of the excess return for each asset. As mentioned above, it is difficult to estimate derivatives precisely using nonparametric density estimation. Therefore, instead of following this route, one could avoid directly estimating the partial derivatives, V_r and V_σ, by considering the instantaneous *covariances* between the asset return and changes in the interest rate/volatility, c_{Vr} and $c_{V\sigma}$. From equations (14), (15) and (16) (after a little simplification),

$$\begin{pmatrix} c_{Vr} \\ c_{V\sigma} \end{pmatrix} \equiv \begin{pmatrix} dV\, dr/V\, dt \\ dV\, d\sigma/V\, dt \end{pmatrix} = \begin{pmatrix} \sigma^2 & \rho\sigma s \\ \rho\sigma s & s^2 \end{pmatrix} \begin{pmatrix} V_r/V \\ V_\sigma/V \end{pmatrix}. \tag{20}$$

This can be inverted, as long as $|\rho| < 1$, to obtain

$$\begin{aligned} \begin{pmatrix} V_r/V \\ V_\sigma/V \end{pmatrix} &= \begin{pmatrix} \sigma^2 & \rho\sigma s \\ \rho\sigma s & s^2 \end{pmatrix}^{-1} \begin{pmatrix} c_{Vr} \\ c_{V\sigma} \end{pmatrix}, \\ &= \frac{1}{1-\rho^2} \begin{pmatrix} 1/\sigma^2 & -\rho/\sigma s \\ -\rho/\sigma s & 1/s^2 \end{pmatrix} \begin{pmatrix} c_{Vr} \\ c_{V\sigma} \end{pmatrix}. \end{aligned}$$

To preclude arbitrage, the excess return on the asset must also be expressible as a linear

combination of c_{Vr} and $c_{V\sigma}$,

$$m = r - \frac{d}{V} + \lambda^*{}_r(r,\sigma)c_{Vr} + \lambda^*{}_\sigma(r,\sigma)c_{V\sigma}. \tag{21}$$

Given two different interest rate dependent assets, we can estimate the instantaneous covariances for each in the same way as we estimated $\rho(r,\sigma)$ above. We can also estimate the excess return for each asset, $m_i(r,\sigma) - r$ as a function of the two state variables. The two excess return can be expressed in the form

$$\begin{pmatrix} m_1 - r \\ m_2 - r \end{pmatrix} = \begin{pmatrix} c^1_{Vr} & c^1_{V\sigma} \\ c^2_{Vr} & c^2_{V\sigma} \end{pmatrix} \begin{pmatrix} \lambda^*{}_r \\ \lambda^*{}_\sigma \end{pmatrix},$$

which can be inverted to yield an estimate of the prices of risk,

$$\begin{pmatrix} \lambda^*{}_r \\ \lambda^*{}_\sigma \end{pmatrix} = \begin{pmatrix} c^1_{Vr} & c^1_{V\sigma} \\ c^2_{Vr} & c^2_{V\sigma} \end{pmatrix}^{-1} \begin{pmatrix} m_1 - r \\ m_2 - r \end{pmatrix}.$$

Finally, for estimates of the more standard representation of the prices of risk, λ_r and λ_σ, equate equations (18) and (21), using equation (20), to obtain

$$\begin{pmatrix} \lambda_r \\ \lambda_\sigma \end{pmatrix} = \begin{pmatrix} \sigma^2 & \rho\sigma s \\ \rho\sigma s & s^2 \end{pmatrix} \begin{pmatrix} \lambda^*{}_r \\ \lambda^*{}_\sigma \end{pmatrix}.$$

Given estimates for the process governing movements in r and σ, and the above procedure for the functions λ_r and λ_σ, we can value interest rate dependent assets in one of two ways. The first is to solve equation (19) numerically using a method such as the Hopscotch method of Gourlay and McKee (1977). The second is to use the fact that we can write the solution to equation (19) in the form of an expectation. Specifically, we can write V, the value of an asset which pays out cash flows at a (possibly path dependent) rate C_t, in the form

$$V_t = E\left[\int_t^T e^{-\int_t^s (\widehat{r}_u)\,du} C_s\,ds\right], \tag{22}$$

where $\widehat{r}$ follows the "risk adjusted" process,

$$d\widehat{r}_\tau = [\mu_r(\widehat{r}_\tau,\widehat{\sigma}_\tau) - \lambda_r(\widehat{r}_\tau,\widehat{\sigma}_\tau)]\,d\tau + \widehat{\sigma}_\tau\,dZ_1, \tag{23}$$

$$d\widehat{\sigma}_\tau = [\mu_\sigma(\widehat{r}_\tau,\widehat{\sigma}_\tau) - \lambda_\sigma(\widehat{r}_\tau,\widehat{\sigma}_\tau)]\,d\tau + \rho s(\widehat{r}_t au,\widehat{\sigma}_\tau)\,dZ_1 + \sqrt{1-\rho^2}s\,dZ_2, \tag{24}$$

for all $\tau > t$, and where

$$\begin{aligned} \widehat{r}_t &= r_t, \\ \widehat{\sigma}_t &= \sigma_t. \end{aligned}$$

This says that the value of the asset equals the expected sum of discounted cash flows paid over the life of the asset, except that it substitutes the risk adjusted process $(\widehat{r}, \widehat{\sigma})$ for the true process (r, σ).

This representation leads directly to a valuation algorithm based on Monte Carlo simulation. For a given starting value of (r_t, σ_t), simulate a number of paths for $\widehat{r}$ and $\widehat{\sigma}$ using equations (23) and (24). Along each path, calculate the cash flows C_t, and discount these back along the path followed by the instantaneous riskless rate, $\widehat{r}_t$. The average of the sum of these values taken over all simulated paths is an approximation to the expectation in equation (22), and hence to the security value, V_t. The more paths simulated, the closer the approximation.

5 Conclusion

This paper provides a method for estimating multifactor continuous-time Markov processes. Using Milshtein's (1978) approximation schemes for writing expectations of functions of the sample path of stochastic differential equations in terms of the drift, volatility and correlation coefficients, we provide non-parametric estimation of the drift and diffusion functions of multivariate stochastic differential equations. We apply this technique to the short- and long-end of the term structure for a general two-factor, continuous-time diffusion process for interest rates. In estimating this process, several results emerge. First, the volatility of interest rates is increasing in the level of interest rates, only for sharply, upward sloping term structures. Thus, the result of previous studies, suggesting an almost exponential relation between interest rate volatility and levels, is due to the term structure on average being upward sloping, and is not a general result per se. Second, the finding that partly motivates this paper, i.e., the link between slope and interest rate volatility in Engle, Lilien and Robins (1987), comes out quite naturally from the estimation. Finally, the slope of the term structure, on its own, plays a large role in determining the magnitude of the diffusion coefficient. These volatility results hold across maturities, which suggests that a low dimensional system (with nonlinear effects) may be enough to explain the term structure

of interest rates.

As a final comment, there are several advantages of the procedure adopted in this paper. First, there is a constant debate between researchers on the relative benefits of using equilibrium versus arbitrage-free models. Here, we circumvent this issue by using actual data to give us the process and corresponding prices of risk. Since the real world coincides with the intersection of equilibrium and arbitrage-free models, our model is automatically consistent. Of course, in a small sample, statistical error will produce estimated functional forms that do not conform. This problem, however, is true of all empirical work. Second, we show how our procedure for estimating the underlying multifactor continuous-time diffusion process can be used to generate fixed income pricing. As an example, we show how our results can be interpreted within a generalized Longstaff and Schwartz (1992) framework, that is, one in which the drift and diffusion coefficients of the instantaneous interest rate and volatility are both (nonlinear) functions of the level of interest rates and the volatility. Third, and perhaps most important, the pricing of fixed-income derivatives depends crucially on the level of volatility. The results in this paper suggest that volatility depends on both the level and slope of the term structure, and therefore contains insights into the eventual pricing of derivatives.

References

Ahn, D., Dittmar, R. and A. Gallant, 2002, Quadratic Term Structure Models: Theory and Evidence, *Review of Financial Studies* 16, 459-485.

Aït-Sahalia, Yacine, 1996a, "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica* 64, 527–560.

Aït-Sahalia, Yacine, 1996b, "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies* 9, 385-426.

Aït-Sahalia, Yacine, 2002, "Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach," *Econometrica* 70, 223-262.

Aït-Sahalia, Yacine, 2007a, "Estimating Continuous-Time Models Using Discretely Sampled Data," Econometric Society World Congress Invited Lecture, in *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*, edited by Richard Blundell, Persson Torsten and Whitney K. Newey, Econometric Society Monographs, Cambridge University Press.

Aït-Sahalia, Yacine, 2007b, "Closed-Form Likelihood Expansions for Multivariate Diffusions," *Annals of Statistics*, forthcoming.

Aït-Sahalia, Yacine and R. Kimmel, 2007a, "Maximum Likelihood Estimation of Stochastic Volatility Models," *Journal of Financial Economics* 83, 413-452.

Aït-Sahalia, Y. and R. Kimmel, 2007b, "Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansions," working paper, Princeton University.

Andersen, T. and J. Lund, 1997, "Estimating Continuous Time Stochastic Volatility Models of the Short Term Interest Rate," *Journal of Econometrics* 77, 343–377.

Bandi, Federico, 2002, "Short-Term Interest Rate Dynamics: A Spatial Approach," *Journal of Financial Economics* 65, 73-110.

Boudoukh, Jacob, Matthew Richardson, Tom Smith and Robert F. Whitelaw, 1999a, "Ex Ante Bond Returns and the Liquidity Preference Hypothesis," *Journal of Finance* 54, 1153-1167.

Boudoukh, Jacob, Matthew Richardson, Tom Smith and Robert F. Whitelaw, 1999b, "Bond Returns and Regime Shifts," working paper, NYU.

Brennan, Michael and Eduardo Schwartz, 1979, "A Continuous Time Approach to the Pricing of Bonds," *Journal of Banking and Finance* 3, 133-155.

Brenner, Robin, Richard Harjes and Kenneth Kroner, 1996, "Another Look at Models of the Short-Term Interest Rate," *Journal of Financial and Quantitative Analysis* 31, 85-107.

Brown, Stephen and Phil Dybvig, 1986, "The Empirical Implications of the Cox, Ingersoll Ross Theory of the Term Structure of Interest Rates," *Journal of Finance* 41, 617-630..

Campbell, John and Robert Shiller, 1991, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies* 58, 495-514.

Chan, K.C., A. Karolyi, F. Longstaff and A. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance* 47, 1209-1227.

Chapman, D. A. and N. D. Pearson, 2000, "Is the Short Rate Drift Actually Nonlinear?" *Journal of Finance* 55, 355-388.

Chen, Ren-Raw and Louis Scott, 1992, "Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure," *Review of Financial Studies* 5, 613-636.

Collin-Dufresne, Pierre, Robert Goldstein and Christopher Jones, 2006, "Can Interest Rate Volatility be Extracted from the Cross Section of Bond Yields? An Investigation of Unspanned Stochastic Volatility," working paper, U.C. Berkeley.

Conley, Timothy, Lars Hansen, Erzo Luttmer and Jose Scheinkman, 1995, "Estimating Subordinated Diffusions from Discrete Time Data," *The Review of Financial Studies* 10, 525-577.

Cox, John C., Jonathan E. Ingersoll, Jr. and Stephen A. Ross, 1985, "A Theory Of The Term Structure Of Interest Rates," *Econometrica* 53, 385-408.

Dai, Qiang and Kenneth J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance* 55, 1943-1978.

Duffie, Darrell, 1988, *Security Markets: Stochastic Models*, Academic Press, Boston.

Duffie, Darrell and Rui Kan, 1996, "A Yield-Factor Model of Interest Rates," *Mathematical Finance* 6, 379-406.

Duffie, Darrell, J. Ma and J. Yong, 1995, "Black's Consol Rate Conjecture," *Annals of Applied Probability* 5, 356-382.

Engle, R., D. Lilien and R. Robins, 1987, "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica* 55, 391-407.

Engle, R. and V. Ng, 1993, "Time-Varying Volatility and the Dynamic Behavior of the Term Structure," *Journal of Money, Credit, and Banking* 25, 336-349.

Fama, Eugene F., 1986, "Term Premiums And Default Premiums In Money Markets," *Journal of Financial Economics* 17, 175-196.

Fama, Eugene F. and Robert Bliss, 1987, "The Information in Long Maturity Forward Rates," *American Economic Review* 77, 680-692.

Gibbons, Michael and Krishna Ramaswamy, 1993, "A Test of the Cox, Ingersoll and Ross Model of the Term Structure," *Review of Financial Studies* 6, 619-658.

Gourlay, A.R. and S. McKee, 1977, "The Construction of Hopscotch Methods for Parabolic and Elliptic Equations in Two Space Dimensions with a Mixed Derivative," *Journal of Computational and Applied Mathematics* 3, 201206.

Hansen, Lars and Jose Scheinkman, 1995, "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes," *Econometrica* 63, 767-804.

Heath, David, Robert Jarrow and Andrew Morton, 1992, "Bond Pricing and the Term Structure of Interest Rates," *Econometrica* 60, 77-105.

Hille, E. and R. Phillips, 1957, *Functional Analysis and Semigroups*, American Mathematical Society, Providence, R.I.

Huizinga, John and Frederic S. Mishkin, 1986, "Monetary Policy Regime Shifts and the Unusual Behavior of Real Interest Rates," Carnegie-Rochester Conference on Public Policy, 24, 231-274.

Ingersoll, Jonathan, 1987, *Theory of Financial Decision Making*, Rowman and Littlefield, Totowa, NJ.

Klemkosky, Robert C. and Eugene A. Pilotte, 1992, "Time-Varying Term Premiums On U.S. Treasury Bills And Bonds," *Journal of Monetary Economics* 30, 87-106.

Litterman, Robert and Jose Scheinkman, 1991, "Common Factors Affecting Bond Returns," *Journal of Fixed Income* 1, 54-61.

Lo, Andrew, 1988, "Maximum likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data," *Econometric Theory* 4, 231-247.

Longstaff, Francis and Eduardo Schwartz, 1992, "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *Journal of Finance* 47, 1259-1282.

Milshtein, G.N., 1974, "Approximate Integration of Stochastic Differential Equations," *Theory of Probability and Its Applications* 19, 557-562.

Milshtein, G.N., 1978, "A Method of Second-Order Accuracy Integration for Stochastic Differential Equations," *Theory of Probability and Its Applications* 23, 396-401.

Øksendal, B., 1985, *Stochastic Differential Equations: An Introduction with Applications,* 3rd edition, Springer-Verlag, New York.

Piazzesi, Monika, 2005, "Bond Yields and the Federal Reserve," *Journal of Political Economy* 113, 311-344.

Pearson, Neal and Tong-Shen Sun, 1994, "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll Ross Model," *Journal of Finance* 49, 1279-1304.

Pritsker, M., 1998 "Nonparametric Density Estimation of Tests of Continuous Time Interest Rate Models," *The Review of Financial Studies* 11, 449-487.

Sanders, A.B. and H. Unal, 1988, "On the Intertemporal Stability of the Short Term Rate of Interest," *Journal of Financial and Quantitative Analysis* 23, 417-423.

Schaefer, Stephen and Eduardo Schwartz, 1984, "A Two-Factor Model of the Term Structure: An Approximate Analytical Solution," *Journal of Financial and Quantitative Analysis* 19, 413-424.

Stambaugh, Robert F., 1988, "The Information In Forward Rates: Implications For Models Of The Term Structure," *Journal of Financial Economics* 21, 41-70.

Stanton, Richard, 1997, "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance* 52, 1973-2002.

Torous, Walter and Clifford Ball, 1995, "Regime Shifts in Short Term Riskless Interest Rates," working paper, London Business School.

Table 1: Conditional Moments of Daily Interest Rate Changes (Basis Points)

The table presents summary statistics for daily changes in the 6-month, 1-year, 3-year, and 5-year yields on U.S. government securities over the 1983 to 2006 period. Specifically, the table provides the mean, volatility, and cross-correlation of these series, conditional on whether the level of the short rate and slope of the term structure are either low or high (and the associated standard errors). These states of the world are labeled HR and LR for high and low short rates, respectively, and HS and LS for high and low slopes, respectively, and they occur with the probabilities given in the first row of the table. A Wald test that the conditional moments are equal (and the associated p-value), holding the short rate state fixed but varying the state for the slope of the term structure, is also provided for the mean and volatility of these series.

	HR,HS	HR,LS	$\chi^2_{HR,HS=HR,LS}$	LR,HS	LR,LS	$\chi^2_{LR,HS=LR,LS}$
Probability	22.76%	26.83%		27.45%	22.96%	
	Mean (bp/day)					
6-month	0.032	-0.292	1.747	0.031	0.056	0.033
(s.e.)/[p-value]	(0.207)	(0.131)	[0.186]	(0.092)	(0.108)	[0.857]
1-year	0.032	-0.365	2.339	0.060	0.069	0.003
(s.e.)/[p-value]	(0.215)	(0.147)	[0.126]	(0.120)	(0.119)	[0.957]
3-year	0.032	-0.365	1.462	0.063	0.082	0.007
(s.e.)/[p-value]	(0.211)	(0.158)	[0.227]	(0.167)	(0.149)	[0.932]
5-year	-0.070	-0.371	1.304	0.017	0.110	0.170
(s.e.)/[p-value]	(0.210)	(0.158)	[0.254]	(0.167)	(0.149)	[0.680]
	Volatility (bp/day)					
6-month	7.645	5.248	35.314	3.715	4.006	0.862
(s.e.)/[p-value]	(0.364)	(0.165)	[0.000]	(0.163)	(0.265)	[0.353]
1-year	7.928	5.869	24.452	4.879	4.428	2.024
(s.e.)/[p-value]	(0.367)	(0.187)	[0.000]	(0.168)	(0.266)	[0.155]
3-year	7.928	5.869	13.564	6.784	5.520	18.173
(s.e.)/[p-value]	(0.341)	(0.187)	[0.000]	(0.180)	(0.229)	[0.000]
5-year	7.746	6.347	13.567	6.761	5.571	20.389
(s.e.)/[p-value]	(0.329)	(0.179)	[0.000]	(0.180)	(0.229)	[0.000]
	Average Correlation					
	0.840	0.823		0.807	0.796	

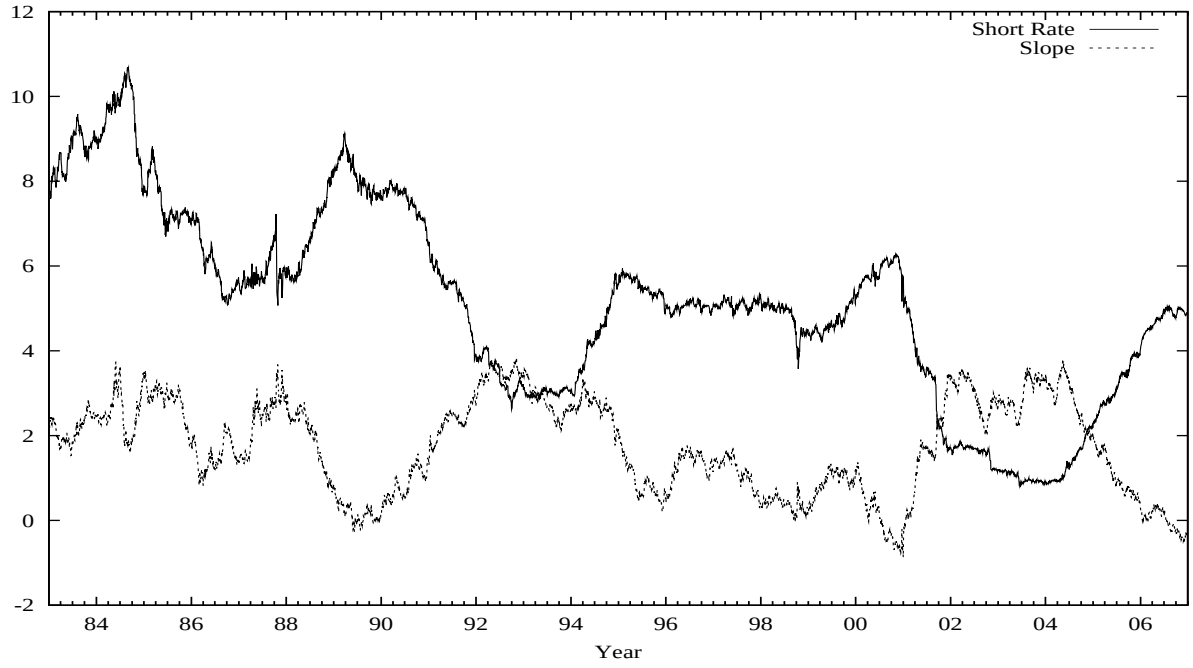

Figure 1: Time series plot of the 3-month rate and term structure slope (i.e., the spread between the 10-year and 3-month rate) over the 1983–2006 period.

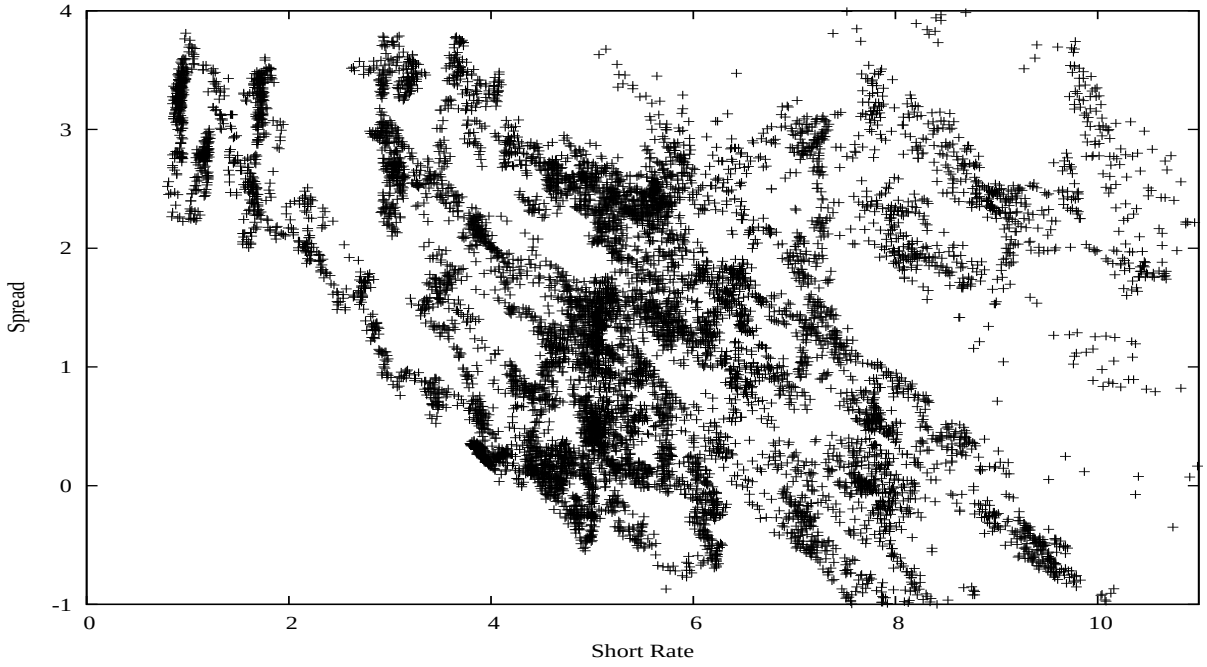

Figure 2: Scatter plot of the 3-month rate vs. the term structure slope over the 1983–2006 period.

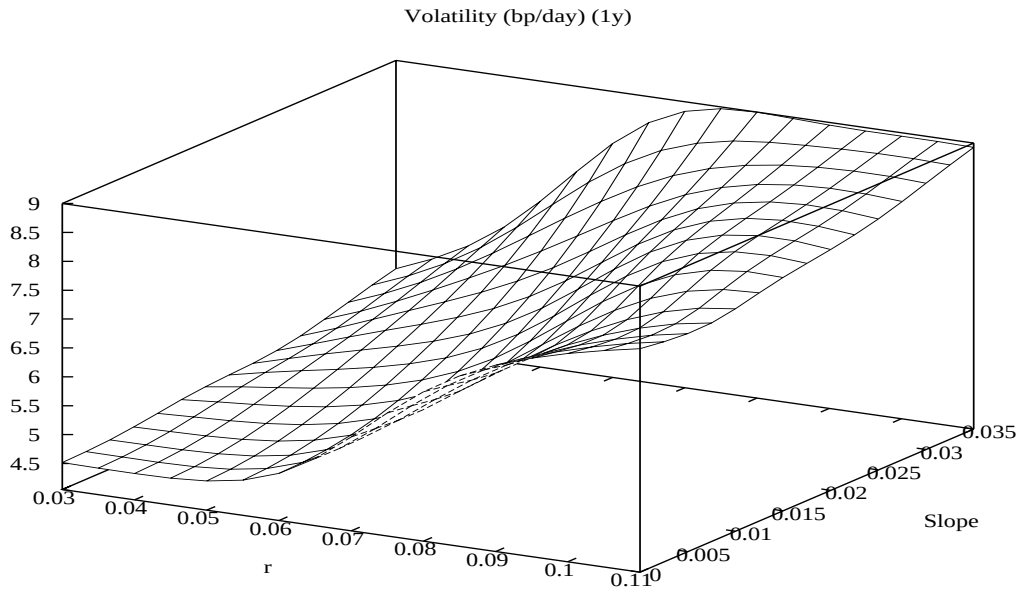

Figure 3: The volatility of the daily change in the 1-year yield (in basis points), conditional on the short rate and the slope of term structure.

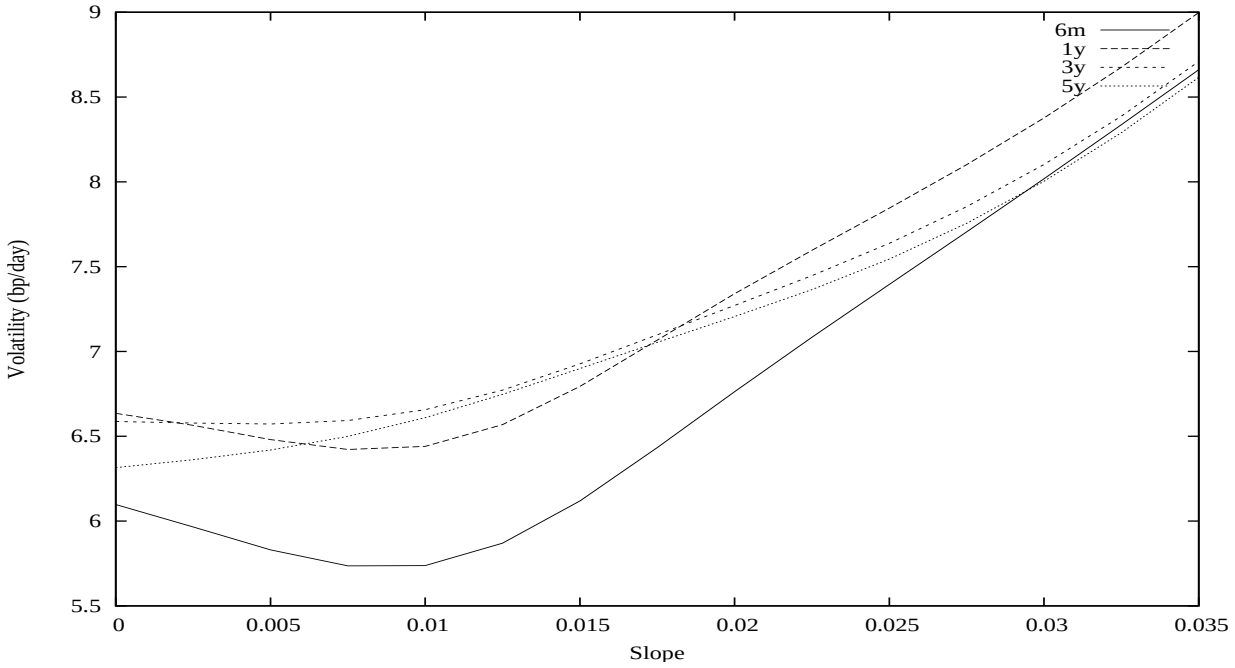

Figure 4: The volatility of the daily change in yields vs. the spread, with the short rate fixed at 8%.

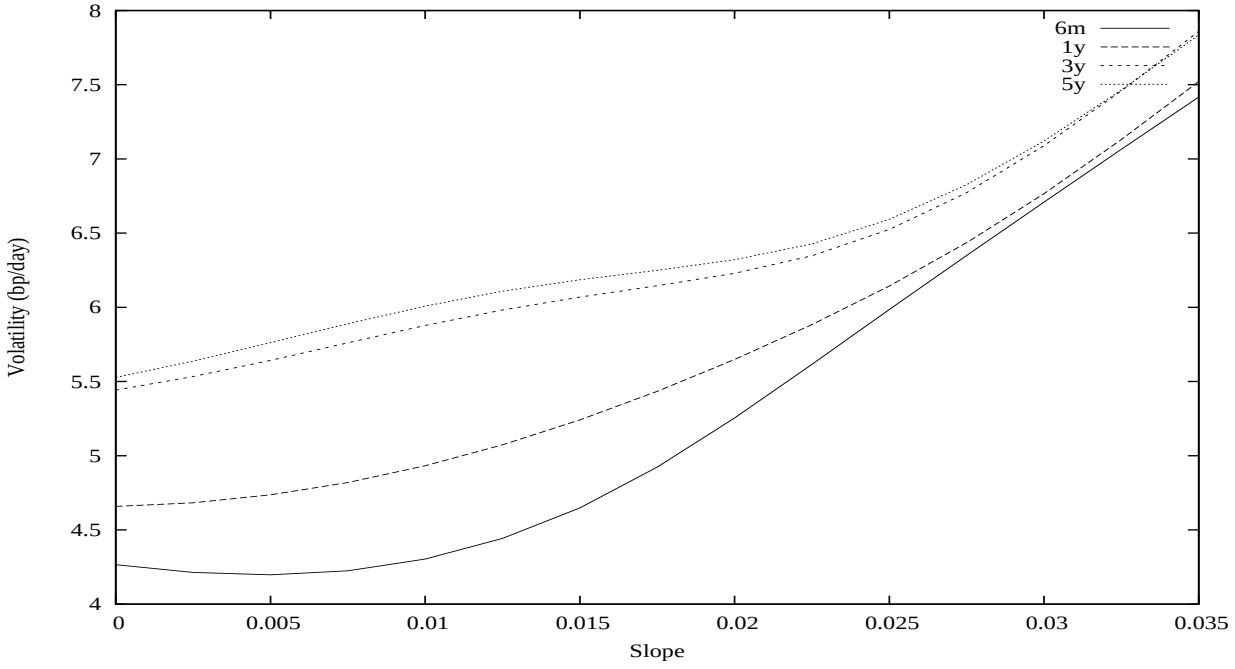

Figure 5: The volatility of the daily change in yields vs. the spread, with the short rate fixed at 5.5%.

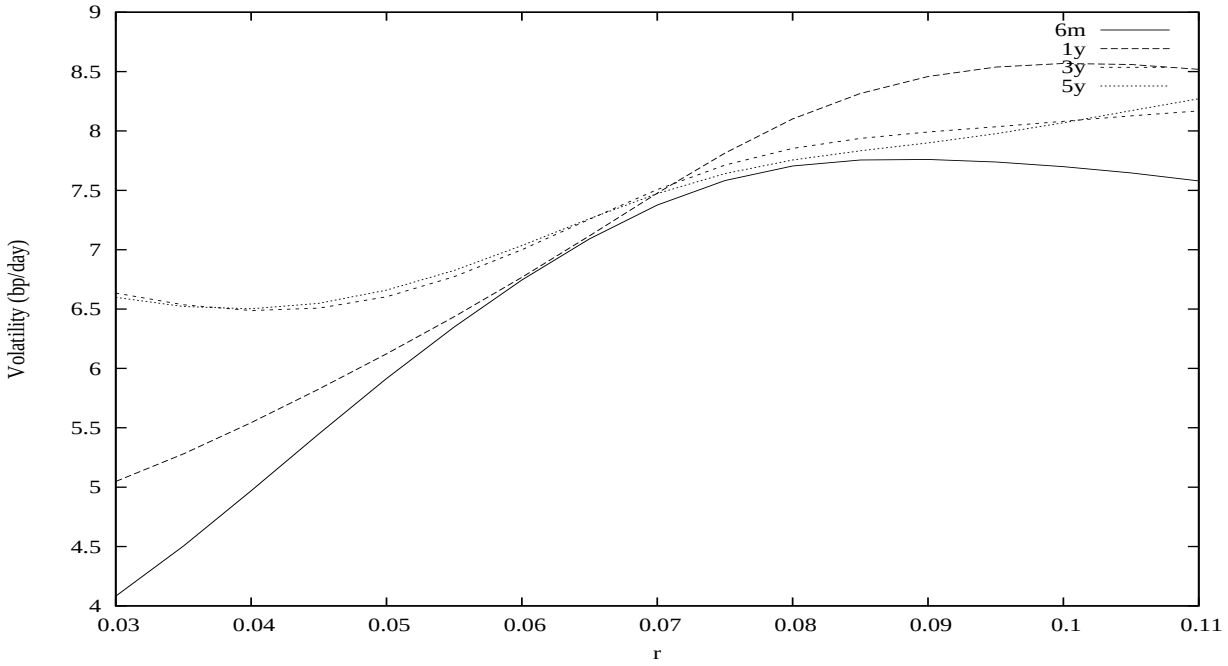

Figure 6: The volatility of the daily change in yields vs. the short rate, with the spread fixed at 2.75%.

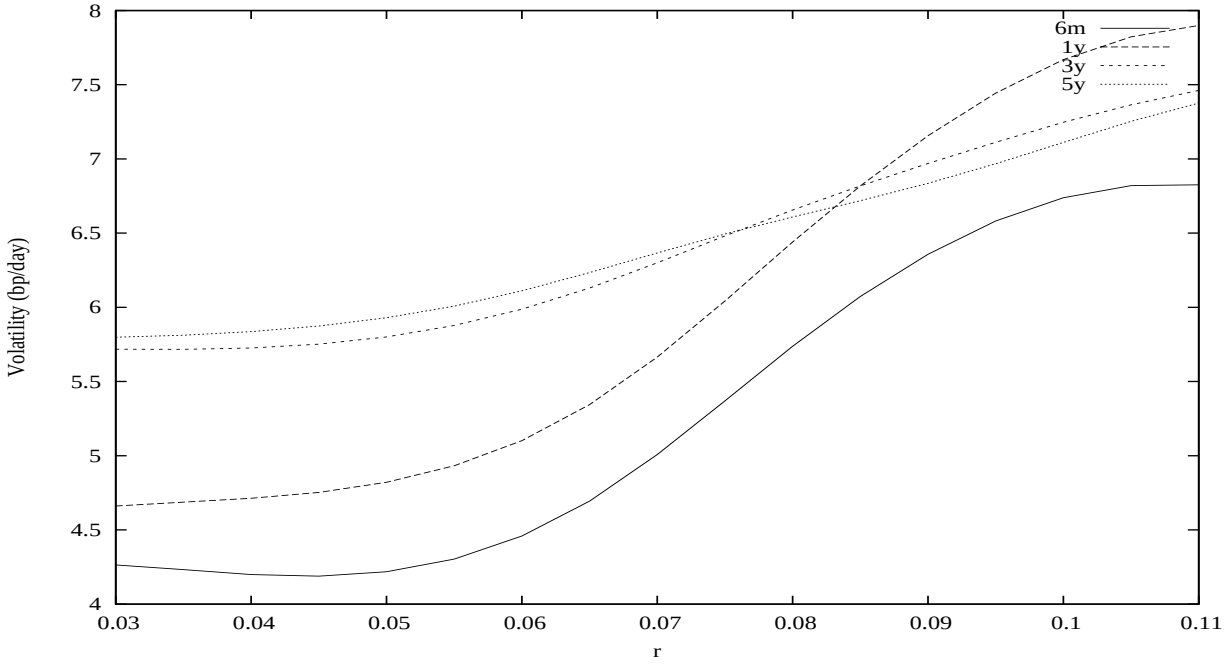

Figure 7: The volatility of the daily change in yields vs. the short rate, with the spread fixed at 1%.

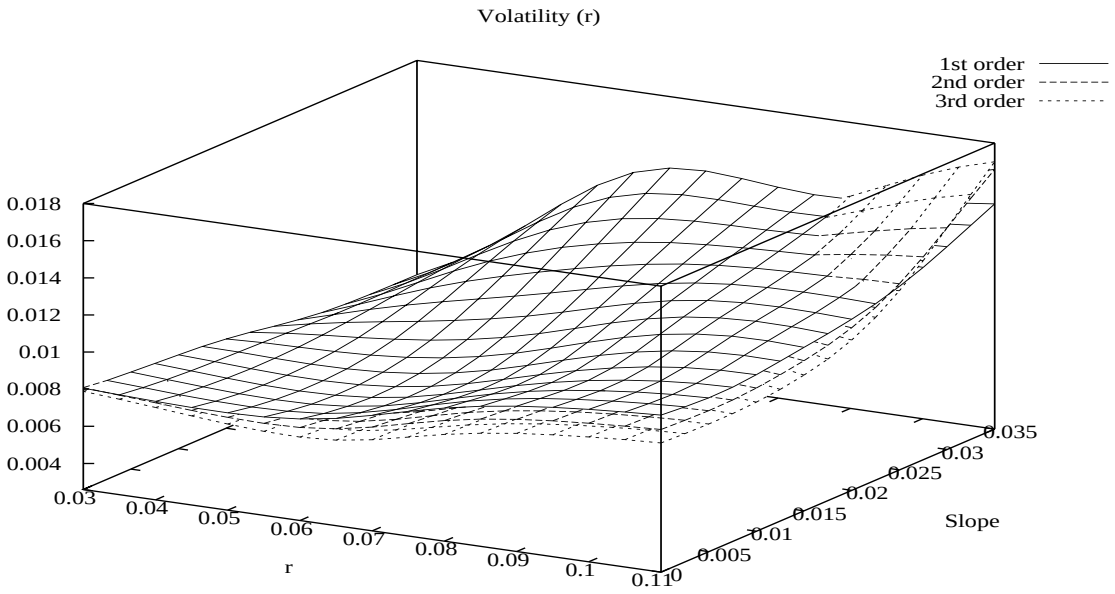

Figure 8: First, second and third order approximations to the diffusion (annualized) of the short rate vs. the short rate and the slope of the term structure.

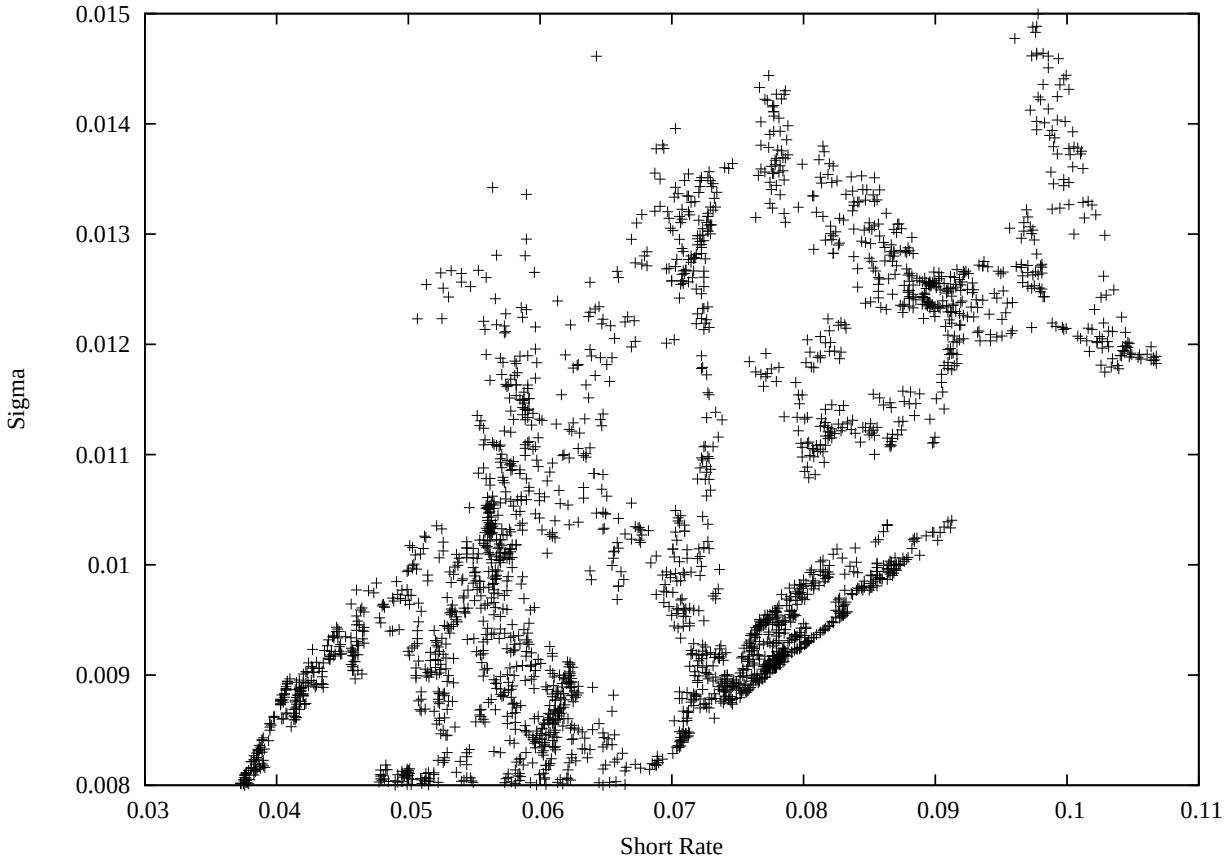

Figure 9: Scatter plot of the 3-month rate vs. the term structure volatility over the 1983–2006 period.

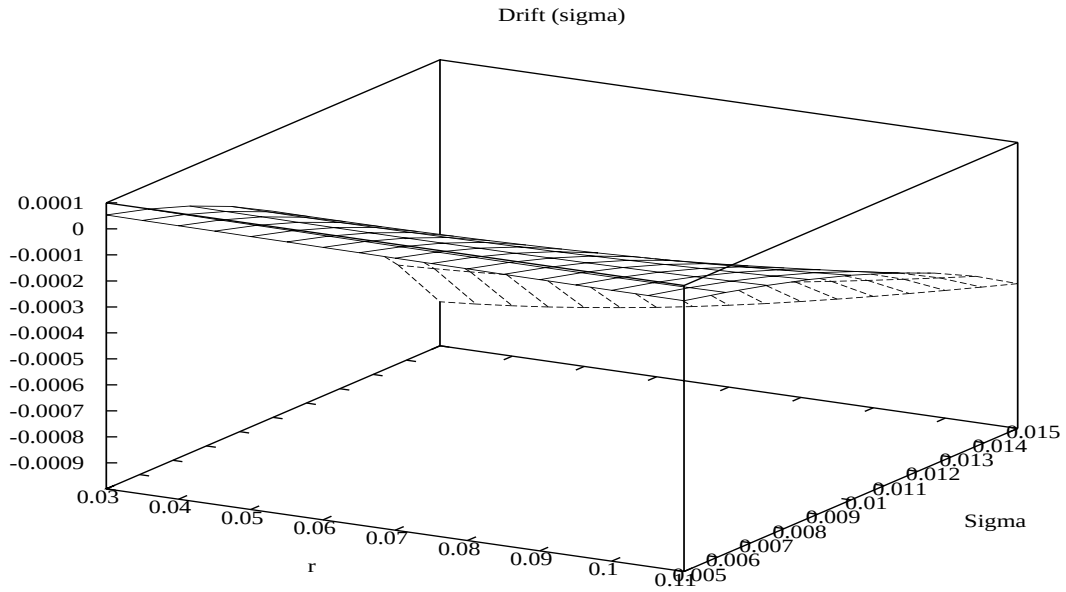

Figure 10: First order approximation to the drift (annualized) of the volatility vs. the short rate and the volatility of the term structure.

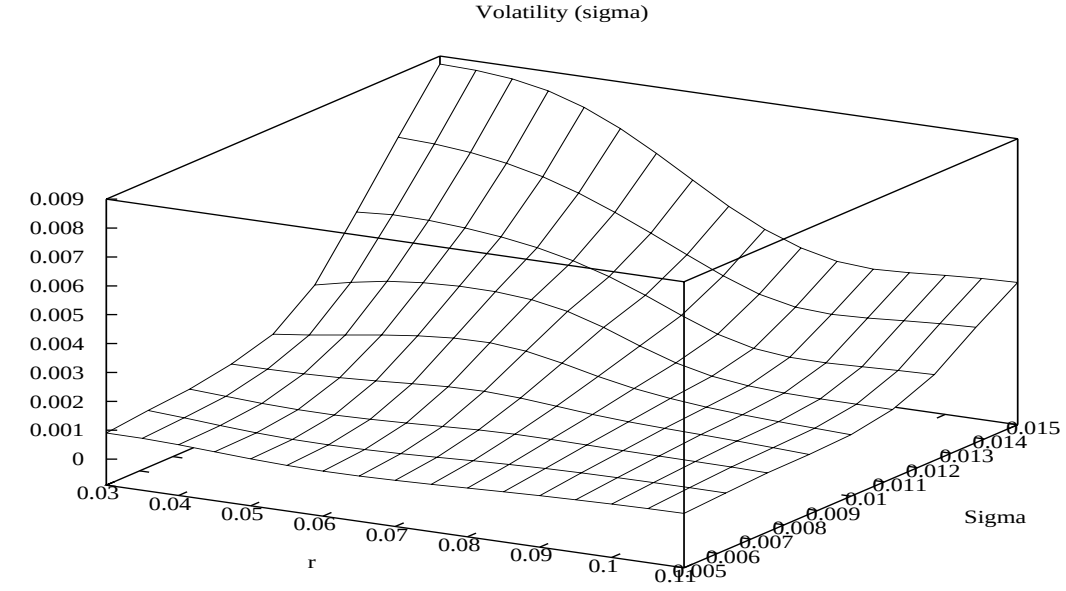

Figure 11: First order approximations to the diffusion (annualized) of the volatility vs. the short rate and the volatility of the term structure.

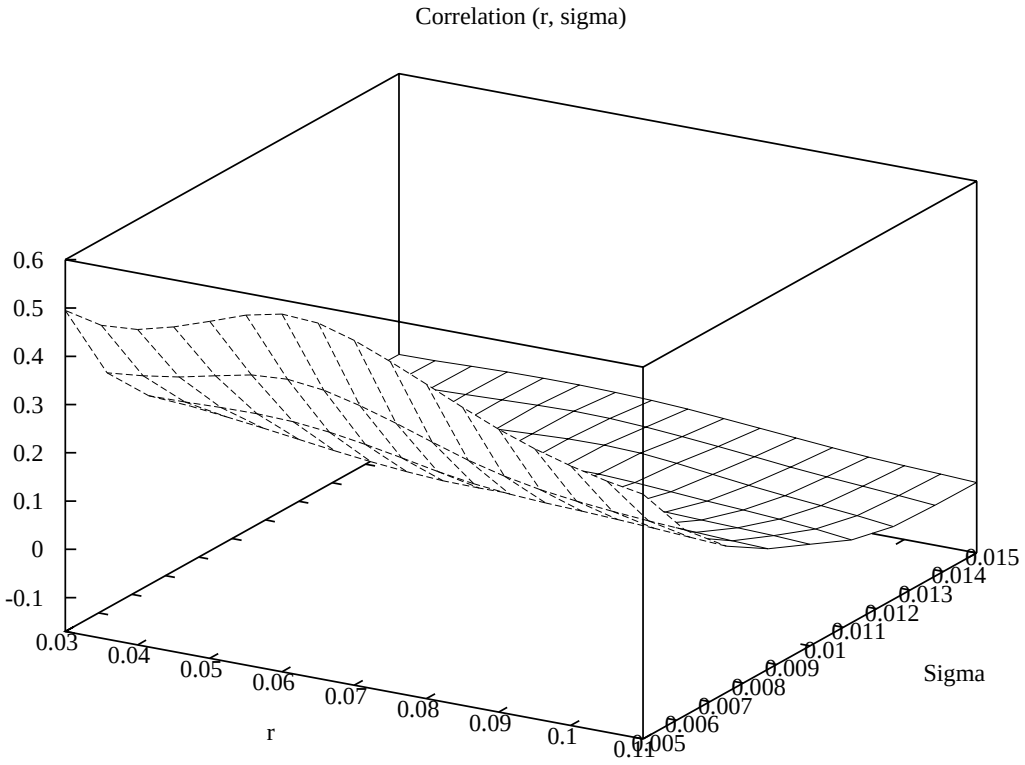

Figure 12: First order approximation to the correlation coefficient between changes in the short rate and the volatility vs. the short rate and the volatility of the term structure.