Time Deformation and the Yield Curve *

Mark Salmon and Wing Wah Tham  
Financial Econometrics Research Centre  
Warwick Business School  

First version: June 2006  
This version: September 2007

Abstract

This paper considers how trading activity at one maturity of the yield curve affects and is affected by trading at other maturities. We approach the modelling of bond prices from a stochastic volatility perspective based on time deformation. We put forward a new, continuous time, multivariate time deformation model which is coherent with the market microstructure theory of price discovery and captures information flow in the market. We model the stochastic volatility process by estimating the instantaneous volatility as a function of the price intensity in the spirit of Cho and Frees (1988) and Gerhard and Hautsch (2002). The point process model, a Hawkes model, that we use to model the price intensity allows for both the self and cross excitation effects of trading at different maturities. Univariate and multivariate models are estimated using transaction level data from BrokerTec, a highly liquid and widely traded electronic platform for US securities.

We find that the integrated price intensity is statistically supported as an appropriate directing process in the US bond market suggesting that private information is revealed indirectly through trades in the presence of information asymmetry and heterogeneous agents. We also find that the individual yields on US treasury notes and bonds appear to be driven by different ‘market clocks’ as suggested perhaps by the market segmentation theory of the term structure. These separate market time scales are then related to each other through a multivariate Hawkes model which effectively coordinates activity along the yield curve. We also show that bond returns standardized by the instantaneous volatility estimated from our Hawkes model are Gaussian which is consistent with the theory of time deformation for security prices quite generally.

JEL: C16, E43, F3, G1, G12  
Keywords: Term structure, Interest rates, Multivariate modeling, Hawkes process, Time deformation.

1 Introduction

Financial markets evolve on a time scale that is invariably different from chronological or clock time which is, after all, only determined by the rate by which the earth revolves around the sun. The relevance of this time scale, which may be natural for the analysis of physical phenomena, can be questioned when fixed intervals of clock time contain differing quantities of information - markets move fast and slow as noted for instance by Hasbrouck (1999). If the underlying stochastic process that we wish to model is indexed to a different time scale than clock time then it becomes difficult, if

*Contact details: Mark.Salmon@wbs.ac.uk, and Wing.Tham03@phd.wbs.ac.uk: Tel +00442476 574168 - We are grateful to Adrian Baddeley, Christian Gourieroux, Nick Webber, Roel Oomen, Nikolaus Hautsch, Hashem Pesaran, Jeffrey Russell, Neil Shephard, Kevin Sheppard, Clive Bowsher, Andrew Patton and Jeremy Large for comments on earlier versions of this paper which were presented at an ESF Exploratory Workshop on The High Frequency Analysis of Foreign Exchange and Fixed Income Markets at Warwick in June 2006, the Econometrics Seminar, Cambridge, the London-Oxford Financial Econometrics workshop at Nuffield, March 2007, Imperial College Financial Econometrics conference London, May 2007 and China International Conference in Finance 2007 at Chengdu, July 2007.
not impossible, to properly measure the properties of these processes without taking time deformation into account.\footnote{Time deformation is the translation of clock time into a time scale that reflects market activity or information flow.} This seems to be crucial for the statistical analysis of security prices quite generally.

Most empirical models of security prices have ignored the question of what time scale should be employed to measure the properties of the asset of interest. Common practice uses a fixed interval of clock time, be it one minute or one day, and these prices do not necessarily reflect the varying amounts of information on the underlying process in these fixed intervals. Although such temporal aggregation superficially facilitates empirical analysis, it masks the true stochastic process that drives asset prices and confuses the ability of statistical methods to measure the process. Hasbrouck (1999) notes that time aggregation of data smears the impact of individual events and aggravates problems of simultaneity. Ait-Sahalia and Mykland (2003) highlights the biases and adverse effects of sampling discreteness and randomness when using a discrete sampling scheme. Information on the underlying process is simply lost on aggregation to fixed intervals of clock time and \textit{ad hoc} assumptions must be used to find representative values for the interval of clock time; such as the choice to use daily closing values to represent the entire day.

Time deformation is a more subtle and deeper issue that is independent of temporal aggregation and rests on the recognition that stochastic processes may evolve on different time scales than natural clock time. For instance consider a sequence of transactions that occur at irregular points in clock time carried out at different prices and with different volume. We can create different time scales defined by volume or price so that the volume time scale is increased by one unit each time, say, 10 million dollars had been transacted since the last “time” point or alternatively the price change had increased by 5 ticks.\footnote{Another classic example is how to value two identical second hand cars that are of the same age but one has done 100,000 miles and the other 10,000. Clearly their common age measured by clock time would not deliver the correct valuation and some account, a time deformation, would have to be made for milage. Second hand car dealers are arbitrageurs in time deformation.} These two time scales would not necessarily coincide of course and represent just two possible choices from a large number of different potential time scales that could be considered; see Le Fol and Mercier (1998). Clark (1973) demonstrates that the subordinated stochastic structure implied by time deformation can potentially reconcile the observed non-Gaussianity of returns, if a proper measure of information flow can be recovered. The argument being that information flow measures the correct time scale on which the underlying price process should be measured and once returns are conditioned on this measure of information flow (or in this time scale) then the central limit theorem applies and Gaussianity is obtained.

This paper considers these issues in the context of understanding the dynamics of the yield curve and addresses the following questions: "Do yields on the different maturities operate on the same time scale, and if not how are they coordinated?", "How is information transferred along the yield curve if different information is relevant at different points on the yield curve?", and "Does this have any implications for the theory of term structure of interest rates as a whole?".

While most models in the time deformation literature are set and estimated in a discrete time framework using data measured in regular intervals, our model and estimation is based on a continuous time framework. We propose a new approach by specifying a structural, continuous time stochastic volatility model for bond prices through a price intensity-based measure of instantaneous volatility. Our volatility estimate is based on how quickly prices change rather than by how much they change, following Cho and Frees (1988). Our approach recognizes the influence of information flow on the market via trading activities and that this information flow is lumpy and non smooth in clock time. Hence it is natural to consider a new time scale which is based on information flow when modeling the statistical properties of bond prices. This is done by using a “frequency” measure of information flow based on the inter-trade transaction times rather than “spatial” measures like volume, order flow or the number of trades. This use of a frequency-based measure of volatility and information flow as a directing process is novel in the literature and provides a convenient extension.
to a multi-variate time deformation model.

In particular, we model the conditional price intensity at different points on the yield curve through a non-homogenous Poisson process that captures the clustering of activity as a function of backward recurrence times and this provides a structural explanation for the transformation between market time and clock time.\(^3\) The estimated price intensity is then used to estimate the instantaneous volatility in the relevant market, in our case, 2, 5 and 30 year US treasury notes and bonds.

The data is modelled as a dynamic point process using transactions level data therefore overcomes the problems associated with econometric models that use regularly spaced data. The model is set in a multi-variate framework accounting for the dependence among price intensities of different maturities in the “frequency” space. Working directly with irregularly spaced data not only circumvents the problem of information loss through data aggregation, it also reduces the possible distortion of the underlying stochastic structure of the original data. This paper is the first, that we know of, to exploit this particular econometric approach to investigate the nature of time-deformation in the treasury market. The analysis is carried out using recently available data from BrokerTec, a highly liquid and widely traded electronic platform for US treasury securities. Thus, this paper opens up new possibilities for understanding and modeling term structure theory as well as the pricing of interest rate related derivatives.

Our results show that US treasury notes and bonds are driven by different information based time scales, but are related to each other through a common ‘market clock’ across the different securities. This finding is consistent with macro finance models which suggest the presence of multiple factors driving the yield curve under the restriction of no-arbitrage.\(^4\) While the macro finance literature finds statistically significant effects of the different direct impacts of macroeconomic news across treasury securities, our results, which measure the indirect impact of information shocks using price intensities, find different price discovery characteristics across securities at different maturities. We can provide two conjectures for this finding. First this could be due to the different impact of news on yields at the different maturities or that different information was relevant at different maturities. Alternatively, this could be explained by different groups of participants with different objectives trading across the securities at the different maturities. Such an interpretation would be closely related to the market segmentation or preferred habitat model of the term structure.\(^5\)

The discovery that Gaussianity is recovered after the return series is standardized by our chronometer supports the view that integrated price intensity is an appropriate directing process for the US bond market.\(^6\) Not only is information partly incorporated directly upon news arrival but it is also revealed indirectly through trades in related assets in the presence of information asymmetry and heterogeneous agents. The significance of our directing process is congruent with the price discovery process in US treasury market found by Green (2004).

In section 2, we discuss our time deformation model and our approach to using a multivariate point process to model stochastic volatility at different maturities on the yield curve. In section 3, we discuss the estimation of our model by Maximum Likelihood Methods and the diagnostic tests we apply. We then describe the data in section 4 and the empirical results in section 5 before drawing some conclusions.

## 2 Time Deformation and Price Intensity

The critical question lies in how to measure information and information flow. A hybrid model based on rational expectations and market microstructure theory suggests that there is both a direct and indirect impact of news on prices. If information and news events are scheduled and anticipated

\(^3\)The backward recurrence time is the time elapsed since the last event and is a left-continuous function before \(t\).
\(^4\)See, for example, Ang and Piazzesi (2003), Duffie and Kan (1996), Dai and Singleton (2000).
\(^5\)See Culbertson (1957) and Modigliani and Sutch (1966) or more recently Vayanos and Vila (2007).
\(^6\)A chronometer is any non-decreasing random process with stationary increments.
by market participants, there will be an instant adjustment of price to this information. This is consistent with the traditional use of the term "price determination" where all information is public and is determined by a consensus combination of the underlying economic fundamentals. In contrast, when a market maker is trading with individuals with private information, his task will be one of discovering the private information. At the heart of many market microstructure models is the fact that uninformed market makers learn about private information from individual orders. Thus, in the presence of information asymmetry and heterogeneous agents, private and heterogeneous information is incorporated indirectly into prices gradually through the process of price discovery\(^7\). It is the measurement of the dynamics of this private information flow process that is critical in modeling time deformation in financial markets.

A number of different proxies for information have been put forward, for example; volume- Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983), Karpoff (1987), Gallant, Rossi and Tauchen (1992), Blume, Easley and O’Hara (1994), the number of trades- Jones, Kaul and Lipson (1994), Ane and Geman (2000), trade duration- Russell and Engle (1998), Engle (2000) and trade intensity- Salmon and McCullough (2005).\(^8\) In this paper, we differ from the literature by using price intensity because of its relationship with the volatility of returns.\(^9\) If informed traders trade more frequently after an information event through a series of price changing trades, there is a direct relationship from private information to the volatility of asset returns via price intensity. This relation allows us to specify a multi-variate stochastic volatility model where the “frequency” based measure of volatility is driven by price intensity in a continuous time framework.

Time deformation has been modeled in several different ways in the literature and a natural representation through stochastic subordination is in terms of stochastic volatility models. Stochastic volatility represents a latent process and one approach to time deformation modeling developed by Stock (1998) assumes that this multivariate latent process \(\mathcal{Z}(s)\) evolves smoothly in operational or market time \(s\). Stock assumes a logistic mapping, \(g(\cdot)\), between market time \(s\) and clock time \(t\): \(s = g(t)\) which describes the relationship between the two time scales (Stock, 1988; Ghysels and Jasiak, 1995). This approach to stochastic volatility modeling is structural if market or economic variables appear within the function \(g(\cdot)\) directing the process by which information flow affects the relationship between clock time and market time and hence whether the market is fast or slow. The approach we follow is similar but employs a different specification for \(g(\cdot)\), by exploiting an underlying point process model for the stochastic price intensity process at each point in the yield curve. The univariate and multivariate Hawkes processes that we describe below allow self and cross excitation from trades at different maturities along the yield curve to account for the clustering of trading activity and information transfers between the different yields. These Hawkes models are then used to model the integrated price intensity which we take as our market time scale and effectively replace Stock’s, ad hoc, logistic specification for \(g(t)\).

Time deformation defines a stochastic volatility model, through a subordinated stochastic process, as originally described by Clark (1973). The directing process in our case, the integrated intensity process, drives stochastic volatility and hence returns. More formally, we model the relation between standard calendar time and market time to be random and the resulting stochastic process is called a chronometer. A subordinator is a special case of a chronometer which has independent and stationary increments.

Asset prices are treated as a semi-martingale:
\[
B(t) = \alpha(t) + m(t),
\]
where \(B(t)\) is the log-price, \(\alpha\) is a predictable process with locally bounded variation and \(m\) is a local martingale. The differential of \(\alpha\) and the differential of quadratic variation of \(m\) could be

\(^7\)Price discovery is a dynamic process in which a diverse group of traders and market makers gather, evaluate, and interpret disparate pieces of information; coordinate trading demands; and generate market-clearing prices.

\(^8\)The use of duration and intensity models in time deformation is also briefly mentioned in Hautsch (2004).

\(^9\)Price intensity is the instantaneous rate of occurrence of a price change of a particular size.
interpreted as the conditional mean and variance of returns respectively. A Levy process could be assumed here as it would allow for a wider class of distributions for returns and the possibility to move away from a Gaussian assumption. However, a Levy process would imply that returns are independent and identically distributed through time and this is widely rejected by empirical work in finance. Instead we will build a model with time deformation, constructing a more flexible model by replacing the natural clock with a market time scale based on a chronometer with non-decreasing random paths with stationary increments. This form of mixture model induces non-gaussianity in the return process.

Our model is then based on time changed Brownian motion, although it could also be used within the more general framework of a Levy process:

\[ B(t) = \alpha(t) + w(\tau^*(t)) , \]  

(2)

where \( w \) is standard Brownian motion which is assumed to be independent from \( \alpha \), the mean process, and \( \tau^* \) is the chronometer. The chronometer is build as

\[ \tau^*(t) = \int_0^t \tau(u) \, du , \]  

(3)

where \( \tau \) is a stationary process. \( \alpha(t) \) is modeled as

\[ \alpha(t) = \mu t + \beta \tau^*(t) , \]  

(4)

which allows the mean process to change with the chronometer. \( \beta \) can be seen as the price of risk when the rate of change in prices is high or a skewness parameter capturing the leverage effect on returns. The resulting models are usually called stochastic volatility models and can also be written in the form of a stochastic differential equation (SDE) with:

\[ dB(t) = \alpha(t) + \sigma(t) \, dw(t) \]
\[ \sigma(t) = \sqrt{\tau(u)} , \]  

(5)

where \( \sigma \) is the spot or instantaneous volatility, while \( \tau \) is the corresponding spot variance. One focus and contribution of our work lies in identifying empirical chronometers, that are statistically supported, with serially dependent increments that reflect the flow of private information into prices as captured by the price intensity model.

Economic considerations in Glosten and Milgrom (1985), Diamond and Verrecchia (1987) and Easley and O’Hara (1992) suggest that durations are indicative of informed trading. In particular, when even the existence of an information event is uncertain, Easley and O’Hara (1992) demonstrate that the lack of trade provides a signal to market participants, i.e. there is no information. A shorter duration between consecutive trades is associated with a higher level of informed trading. Dufour and Engle (2000) highlight the crucial role of duration empirically in assessing the price impact of a trade and Gerhard and Hautsch (2002) use various price intensities and compare the different estimates of volatility that result. Drawing on this theoretical and empirical background regarding the information content in prices and trade durations, we build a stochastic volatility model with a chronometer that is based on the price intensity of the assets.

\footnote{A predictable process is a deterministic process or a c\'{a}gl\'{e}d process which is process which is left-continuous with right hand limits.}
In particular, we estimate the volatility process directly from price intensity in the following way:

$$\sigma^2_c(s) = \lim_{\Delta \to 0} \frac{1}{\Delta} \left[ \text{prob} \left( |B(t) - B(t-\Delta)| \geq c \mid \mathcal{F}_t \right) \right] \times \left( \frac{c}{B(t-\Delta)} \right)^2$$  \hspace{1cm} (6)

$$\sigma^2_c(s) = \lim_{\Delta \to 0} \frac{1}{\Delta} \left[ \text{prob} \left( N^c(t+\Delta) - N^c(t) > 0 \mid \mathcal{F}_t \right) \right] \times \left( \frac{c}{B(t-\Delta)} \right)^2$$  \hspace{1cm} (7)

$$\sigma^2_c(s) = \lambda^c(t; \mathcal{F}_t) \left( \frac{c}{B(t-1)} \right)^2$$  \hspace{1cm} (8)

where $$\tau^*(t) = \int_0^t \lambda^c(u) du = \int_0^t \left[ \mu + \int_{(0,t)} W(t-u, \theta) dN(u) \right] du,$$  \hspace{1cm} (9)

where $$\lambda^c(t; \mathcal{F}_t)$$ is the conditional price intensity of a price change of magnitude $$c$$ at time $$t$$, and $$\sigma^2_c(s)$$ is the directing stochastic instantaneous volatility. $$\lambda^c(t; \mathcal{F}_t)$$ can also be seen as the conditional probability per unit time of observing an absolute price change greater than or equal to $$c$$ in the next instant, given the conditioning information of the backward recurrence time, $$W(t-u)$$. Based on these probabilities, we calculate a conditional instantaneous volatility, as the risk of a price change exceeding the magnitude $$c$$ that an individual is exposed to in the next instant. In doing so, we propose a continuous time stochastic volatility model where returns are driven by a chronometer that is motivated by market microstructure theory.

The mapping between market time, $$\tau^*(t) = \int_0^t \lambda^c(t; \mathcal{F}_t)$$, and the calendar time, $$t$$, given by the stochastic intensity model, specifies the structural stochastic volatility model. Such a specification for volatility, where the instantaneous volatility is modeled using the price intensity, is not new to financial econometrics (see for instance, Cho and Frees (1988), Engle and Russell (1998) and Gerhard and Hautsch (2002)). However, there are no applications of such an approach to explain time deformation and stochastic volatility, as far as we know, in the literature.

This form of volatility estimator examines the time required for prices to move beyond a given price interval using the notion of “first passage time”. While the normal volatility estimator focuses on how much the prices changes, the intensity-based estimator focuses on how frequently the price changes. Cho and Frees (1988) argue that this frequency based estimator will asymptotically eliminate the impact of microstructure biases created, for instance, by the bid ask bounce from the resulting estimate of volatility. In addition volatility specified in this way reflects economic intuition linking the stochastic arrival rate of trades to the ‘stylized fact’ of volatility clustering. As new information arrives, if informed traders are impatient and carry out order splitting strategies to exploit their information advantage as quickly as possible, as suggested by the market microstructure literature, there will be a clustering of trades and thus volatility clustering. This also suggests that private information is incorporated into prices through packets of trades, either fast or slow.

Empirically this approach is also supported by the results found by Garbade and Lieber (1977); that a homogenous Poisson process captures most trading activity relatively well, except for irregular bursts of trade arrivals which invalidate the homogeneous assumption, and Engle and Russell (1998) who observe trade clustering in the NYSE. The model we develop below draws in particular on Hawkes (1971) and Bowsher (2005). As opposed to Engle (2000) or Bauwens and Hautsch (2003), we provide a direct specification for the stochastic intensity process in terms of the backward recurrence time of events rather than through a duration specification. The autoregressive conditional duration framework becomes impractical in the multivariate case as it is difficult to account for the events arriving in one process during the duration spell of another.

The next section describes the specification of the conditional intensity, the estimation technique, stationarity conditions for our model and the diagnostic tests we have used.
3 Stochastic Intensity; Estimation and Testing

3.1 Stochastic Intensity Model

First, assume \( N(t) \) to be a simple point process in \([0, \infty)\) on \((\Omega, \mathcal{F}, P)\) that is adapted to some filtration \(\mathcal{F}_t\), and that \(\lambda(t | \mathcal{F}_t)\) is a positive process with a sample path that is right continuous with left limits. Then, the conditional intensity can be written as:

\[
\lambda(t | \mathcal{F}_t) = \lim_{\Delta \to 0} \frac{1}{\Delta} E \left[ N(t + \Delta) - N(t) | \mathcal{F}_t \right],
\]

where \( N(t) \) represents the number of events that have occurred up to and including time \( t \). We refer to \( \lambda(t | \mathcal{F}_t) \) as \( \lambda(t) \), the value of \( \lambda \) at time \( t \), and \( \mathcal{F}_t \) is defined as the natural filtration up to and including time \( t \). Within this framework, we assume that information flow in the market is captured by the price intensity process; the higher the price intensity the shorter price duration indicating the presence of private information.

The stochastic intensities for the different treasury securities can now be fully specified in a multi-variate framework, where each maturity in the yield curve (or the number of variates) is represented by the price intensity process. In the multi-variate case, \( \lambda_m(t, \theta) \) is a function of \( t \) and \( \theta \), where \( \theta \) is a vector of unknown parameters. The functions within \( W \) will be specified in our model as nonlinear autoregressive excitation processes, where the pattern of trading in the past determines current intensities and trading activity. In the multi-variate case, \( \lambda_m(t, \theta) \) is driven by the backward recurrence time of \( m \) autoregressive exciting effects; each capturing the past occurrence of type \( m \) events, 3month, 12 month trades, etc. The diagonal entries of \( W \) will be called self-exciting effects, while off-diagonal entries are cross-exciting effects. Note that if \( \mu_m = 0 \) then the self- and cross exciting effects in \( W \) would never be realized as there would be nothing to start the self exciting process and so \( \lambda \equiv 0 \). For non-trivial results we therefore require that \( \mu_m \) be positive. Previous work specifies \( \mu_m \) using a spline or a fast Fourier transform (i.e. Bowsher, 2005; Hall and Hautsch, 2004) in order to capture the intra-day deterministic effects. Since there is no clear economic reason why a diurnal pattern should exist in price duration, unlike trade duration, as clustering arrivals of price changing trade are related to information flow of different sources which can occur with different rates anytime throughout the day, we have chosen to specify \( \mu_m \) as a vector of constants rather than assume some, potentially misspecified, deterministic diurnal function.

We specify the function \( W \) using Hawkes models for the stochastic intensity with a non-negative exponentially decaying function, which captures the clustering of trading. This allows us to measure the impact of periods with different information flow and their rates of decay. Note that \( \mu_m \) must be non-negative and that different forms of the decay function of the backward recurrence time are possible but, in common with earlier work, we will only employ an exponential form: 

\[
W(t - u, \theta) = \alpha_{m,r}^{j} \exp \left[ -\beta_{m,r}^{j}(t - \tau_{r,k}) \right].
\]

The final specification of the multi-variate stochastic intensities is then given as:

\[
\lambda_m(t) = \mu_m + \sum_{r=1}^{M} \sum_{j=1}^{D} \sum_{k=1}^{N} \alpha_{m,r}^{j} \exp \left[ -\beta_{m,r}^{j}(t - \tau_{r,k}) \right], \tag{12}
\]

\[\]For a more detailed analysis of the asymptotic properties and stationarity conditions see Hawkes (1971).
where $M$ is the number of maturities we consider, $D$ is the number of dimensions (exponential functions), $N_r$ is the number of data points of type $r$, and $\tau_{r,k}$ is the occurrence time for the $k^{th}$ data point of process $m$.

In this model, we, therefore, allow market time to take multiple scales corresponding perhaps to different types of information— if there were a single market time scale then a univariate Hawkes model would be common for all maturities or the impact factor $\alpha_{m,r}^{j}$ will be the same for all $\alpha^{-s}$.$^{12}$ A graphical illustration of a Hawkes model is presented in Figure 1 below which shows the self exciting nature of the model as past trades stimulate the possibility of current trading.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{hawkes_model}
\caption{Graphical Illustration of a Hawkes Model}
\end{figure}

4 Maximum Likelihood Estimation

Ogata (1978) establishes, under certain regularity conditions, that the MLE for a simple, stationary univariate point process model is consistent and asymptotically normal as $T \to \infty$, and the likelihood ratio test of a simple null hypothesis possesses the standard $\chi^2$ asymptotic null distribution.

The assumption that the point process is simple implies that the intensity of each type $m$ gives the conditional probability of a trade occurring per unit of time as the time interval tends to zero. If the sample path of the point process is integrable, then there exists an analytic form for the likelihood function of the specified conditional intensity. Given a parameterized Hawkes’ model, the unknown

$^{12}$Our model differs from the g-HawkesE model in Bowsher (2005), which uses only the past events of the same trading day and the previous day’s intensity. This specification is more flexible, but it is difficult to derive conditions for stationarity, as the additive effect of the previous events is fixed as endpoints of each individual trading day.
parameters, \( \theta = (\mu_m, \alpha_{m,r}, \beta_{m,r}) \), can be estimated using MLE with a likelihood function given by:

\[
\mathcal{L} \left( \theta; \{ N(t) \}_{t \in (0,T]} \right) = \exp \sum_{m=1}^{M} \left[ \int_{(0,T]} \log(\lambda_m(s;\theta | \mathcal{F}_s)) dN_m(s) + \int_{0}^{T} (1 - \lambda_m(s | \mathcal{F}_s)) \, ds \right].
\]  

(13)

The parameters associated with \( \lambda_m(t;\theta | \mathcal{F}_s) \) in a multivariate Hawkes model are often assumed to be variation-free or separable (Bowsher, 2005; Hall and Hautsch, 2004), allowing for separate maximization of \( m \) log likelihood components and avoiding the curse of dimensionality. Separability is important in a point process model because it eases the complexity of estimation as a parameter or a set of parameters may be estimated individually.\(^{13}\) However, as far as we can see, the full impact of the variation-free assumption or separability has not yet been rigorously examined in this particular context and the impact on the asymptotic properties of the MLE estimates are therefore still unclear. We have been able to avoid any assumption of separability by rewriting the joint likelihood in a recursive conditional manner and directly maximizing the joint likelihood following a procedure outlined in Sohrmann and Tham (2006).

4.1 Stationarity Conditions

Assuming \( \mu \) to be constant rather than a spline allows one to derive the stationarity conditions for the multivariate case. If

\[
\lambda_s(t) = \mu_s + \sum_{r=1}^{k} \int_{-\infty}^{t} g_{sr}(t-u) dN_s(u)
\]

(14)

\[
\Lambda(t) = \mu + \int_{-\infty}^{t} G(t-u) dN(u),
\]

(15)

then, the vector of stationary densities is:

\[
\Lambda = (1 - \Gamma)^{-1} \mu
\]

(16)

\[
\Gamma = \int_{0}^{\infty} G(v) dv,
\]

(17)

given \( \Lambda > 0 \). For an invertible \((1 - \Gamma)\), this condition will give \( k \) rows of stationarity conditions. This integral equation is in general difficult to solve analytically. However, an analytic solution may be obtained when \( g(t-v) \) decays exponentially (Hawkes, 1971) and in our case corresponds to the multivariate stationarity condition.

\[
0 < \sum_{r=1}^{M} \sum_{j=1}^{D} \frac{\alpha_{m,r}^{j}}{\beta_{m,r}} < 1
\]

In each case reported below in our empirical results we find this condition is satisfied and therefore the models we have estimated are stationary.

4.2 Generalized Residuals and Specification Testing

A common approach to evaluating point process models is to examine discrimination criteria such as the Akaike Information Criterion or the Bayesian Information Criterion (e.g. Ogata, 1988). These provide useful numerical comparisons of the relative goodness of fit of competing models but cannot

\(^{13}\) For a more detailed discussion of separability in multi-dimensional point process, refer to Schoenberg (2003), (2004).
Residual analysis in other statistical contexts is a standard and powerful tool for locating defects in the fitted model and for suggesting how the model should be improved. While the same is true in point process modeling, one needs to be careful in how you define the residuals in this context as they will themselves form yet another point process. The ensuing residual process is therefore similar to the generalized residuals considered in Cox and Snell (1968) and diagnostic checking in the point process literature is carried out by examining various plots constructed from the residual process, see Baddeley et al (2005).

If \( N(t) \) denotes the count process of trades then residuals can be constructed from the fact that the unobservable error or innovation process;

\[
I(t) = N(t) - \int_0^t \lambda(s)ds,
\]

is a martingale with \( E[I(t)] = 0 \) when the model is true. When the point process model is fitted to the data and parameters, \( \theta \), estimated as, \( \hat{\theta} \), then the estimated parameters could be substituted into \( \lambda(t) = \lambda_0(t) \) to give the estimated conditional intensity \( \hat{\lambda}(t) = \lambda_0(t) \) from which we can compute the raw residual process

\[
R(t) = N(t) - \int_0^t \hat{\lambda}(s)ds.
\]

Increments in \( R(t) \) are equivalent to the residuals (observed minus fitted) in a regression model and the adequacy of the fitted model can be examined by checking if \( R(t) \approx 0 \). Another standard approach is to examine if the estimated unconditional intensity function \( \lambda(t) \) delivers a homogenous Poisson residual process. Suppose we observe a one-dimensional point process \( t_1, \cdots, t_n \) with conditional intensity \( \lambda(t) \) on an interval \([0, T]\). Papangelou (1974) shows that the integrated conditional intensity of the process forms a homogenous Poisson process with a unit rate on the interval \([0, n]\). If the estimated intensity \( \hat{\lambda}(t) \) is close to the true conditional intensity, then the residual process should resemble a unit rate homogenous Poisson process.

**Theorem 1** Let \( N(t) \) be a simple point process on \([0, \infty)\). Suppose that \( N(t) \) has the intensity function \( \lambda(t \mid \mathcal{F}_t) \) that satisfies:

\[
\int_0^\infty \lambda(t \mid \mathcal{F}_t) dt = \infty,
\]

define for \( \forall t \), the stopping time \( \tau_t \) as the solution to:

\[
\int_0^{\tau_t} \lambda(s \mid \mathcal{F}_s) ds = t,
\]

then, the point process \( \tilde{N}(t) = N(\tau_t) \) is a homogenous Poisson process with intensity \( \lambda = 1 \). Proof is shown in Bremaud (1981).

The only condition for the above relations to hold is the assumption of a simple point process where there is zero probability of the occurrence of more than one event at any single point in time. From the theorem, it can be shown that \( t_i - \tilde{t}_{i-1} = \int_{\tilde{t}_{i-1}}^{t_i} \lambda(s \mid \mathcal{F}_s) ds = \Lambda(t_{i-1}, t_i) \), where \( \{\tilde{t}_i\}_{i \in \{1, 2, \ldots\}} \) denotes the time of occurrence for the sequence of points associated with \( \tilde{N}(t) \). Then it follows that:

\[
\Lambda(t_{i-1}, t_i) \sim \text{i.i.d. Exp}(1).
\]

Note that the above transformation is a time-series transformation of a non-homogeneous Poisson process into a homogenous Poisson process. Moreover, \( \Lambda(t_{i-1}, t_i) \) establishes the link between the intensity function and the duration until the next occurrence of an event. \( \Lambda(t_{i-1}, t_i) \) can be seen as a generalized residual and indicates whether the specified intensity function under-predicts \( \Lambda(t_{i-1}, t_i) < 1 \) or over-predicts \( \Lambda(t_{i-1}, t_i) > 1 \) the number of events at any point in time.
5 Data Description

Over the past several years, trading in the U.S. Treasury securities market has migrated from voice-assisted brokers to fully electronic platforms (Mizrach and Neely, 2006) with nearly all interdealer trading now taking place via one of two electronic communications networks, BrokerTec and eSpeed. Our data is drawn from BrokerTec, an inter-dealer electronic trading platform of secondary wholesale U.S. treasury bonds that currently has a market share of approximately 60-65% of the active issues. It functions as a limit order book and operates about 22-23 hours per day during the week although 95% of trading occurs between 7:30 to 17:30 EST (Fleming, 1997). BrokerTec is a fully automated electronic trading platform where buyers are matched to sellers without human intervention. The brokers provide electronic screens which display the best bid and offer prices and associated quantities. The database provides a comprehensive record of every trade and order book change over the BrokerTec system for the on-the-run 2-, 3-, 5-, and 10- year Treasury notes as well as the 30-year Treasury bond. The trade data include price, quantity, and whether a trade was seller-initiated or buyer-initiated. The order book data specifies the price, quantity change, total quantities for that order, whether the order is a bid or an ask, and the reason for the change. Trades and order book changes are time-stamped to the millisecond. Tick sizes on the BrokerTec platform vary by security. Treasury notes and bonds are quoted in 32nds of a point, where a point equals one percent of par, but the 32nds themselves can be split into halves and, for some securities, quarters. For the 2-, 3-, and 5-year notes, the tick size is \( \frac{1}{4} \) of a 32nd (or 1/128th) of a point (or 0.0078125% of par). For the 10-year note and 30-year bond, the tick size is \( \frac{1}{2} \) of a 32nd (or 1/64th) of a point (or 0.015625% of par). In the BrokerTec database, prices are reported in 256ths of a point. Note that the tick size for the 2-, and 5-year notes is 2/256ths and the tick size for the 30-year bond is 4/256ths.

This paper focuses on the on-the-run notes and bonds, namely 2-year, 5-year notes, and the 30-year bond, traded between 7:30 and 17:30 EST. Even though the on-the-run securities represent just a small fraction of all the outstanding treasury securities, they account for about 71% of the activity in the interdealer market (Fabozzi and Fleming, 2000). Our sample period is January 18th, 2005 to December 30th, 2005. Weekends, holidays, and days with unusually low or no trading activity are removed. The period of estimation is chosen from 7:30 to 17:30 EST as this range of time captures most of all trading activity. Because of the time gap between each trading day, the data is concatenated together so that 17:30 on Monday is followed by 7:30 on Tuesday in a new time line.

5.1 Preliminary Statistics and Trend in Price Duration

Figure 2 depicts the typical average daily price duration on BrokerTec of 5-year treasury notes. The trend pattern is similar for the 2-, 5- and 30-year treasury securities. The findings are consistent with what Fleming (1997) found using GovPX data from 1994. Trading peaks between 08:30 and 09:00, and then again between 10:00 and 10:30. Trading reaches a final peak between 14:30 and 15:00 and then tapers off by 17:30. As suggested by Fleming (1997), this pattern is probably largely explained by scheduled macroeconomic announcements (most of which are made at 08:30 and 10:00) and the hours of open outcry Treasury futures trading (08:20 to 15:00). Thus such patterns in price duration are most probably caused by information inflow into the treasury securities through some price discovery process after a scheduled information arrival and unscheduled information spillover from the futures market.

Table 1 shows some preliminary statistics. There are 28644, 116163, and 84651 observations for the 2-year, 5-year notes, and 30-year bond respectively for trades with an absolute price change greater than or equal than 1 tick. About 98%, 95% and 70% of trades in 2, 5 and 30 year treasury securities respectively have an absolute price change of 1 tick. Thus we have chosen to model our intensity-based volatility using thinned trades with an absolute change in price that is greater than
Figure 2: Intraday pattern of Price Duration of 5-year Notes.

<table>
<thead>
<tr>
<th>Variables</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>276.43</td>
<td>69.2</td>
<td>93.26</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>446.18</td>
<td>113.28</td>
<td>138.76</td>
</tr>
<tr>
<td>Skewness</td>
<td>4.28</td>
<td>5.61</td>
<td>4.96</td>
</tr>
<tr>
<td>Min</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Max</td>
<td>8075.01</td>
<td>3510.89</td>
<td>4015.59</td>
</tr>
<tr>
<td>Observations</td>
<td>28644</td>
<td>116163</td>
<td>84651</td>
</tr>
<tr>
<td>% of trade with 1 tick price change</td>
<td>98%</td>
<td>95%</td>
<td>70%</td>
</tr>
</tbody>
</table>
or equal to 1 tick:

\[ \sigma_{c=1 \text{tick}}^2 (s) = \lambda_c(t; \mathcal{F}_t) \left( \frac{1 \text{ tick}}{B(t - 1)} \right)^2. \]  

(21)

The average duration between trades with a price change of 1-tick is 276.43 seconds, 69.2 seconds, and 93.26 seconds respectively. The minimum duration for all securities is 0.1 second while the maximum duration varies between 3510.886 seconds to 4015.596 seconds. Figures 3-5 show the histograms of the price durations for each security, which suggest over-dispersion.

Figure 3: Histogram of the price duration for 2-year notes.

Figure 4: Histogram of the price duration for 5-year notes

The sizable differences between the means and standard deviations also suggest that the durations are not conditionally exponentially distributed. The Kolmogorov-Smirnov test in Table 2 confirms the rejection of the null that price duration is exponentially distributed. This is congruent with Garbade and Lieber’s (1977) finding that a homogenous Poisson process is not appropriate for modeling trade arrivals.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymptotic Statistics</td>
<td>146.86</td>
<td>229.37</td>
<td>211.96</td>
</tr>
<tr>
<td>P-value</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 2: Kolmogorov-Smirnov test of exponential distribution for price duration

Since we propose to use Hawkes models to capture temporally dependent trades, we first examine
the duration between price-changing trades for signs of potential clustering. The real time autocorrelations in Table 3 are all significantly positive. The Ljung-Box statistic of no autocorrelation is a $\chi^2_{15}$ variable with a 5% critical value of 25 and the null of no autocorrelation is therefore easily rejected for all three securities. So, as in Engle and Russell (1997), we find a clear pattern of temporal dependence in price changing trades in the U.S. treasury market.

Table 3: Duration autocorrelation

<table>
<thead>
<tr>
<th>Duration</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag 1</td>
<td>0.19</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>Lag 2</td>
<td>0.19</td>
<td>0.25</td>
<td>0.22</td>
</tr>
<tr>
<td>Lag 3</td>
<td>0.18</td>
<td>0.24</td>
<td>0.22</td>
</tr>
<tr>
<td>Lag 4</td>
<td>0.16</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>Lag 5</td>
<td>0.15</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>Lag 6</td>
<td>0.14</td>
<td>0.20</td>
<td>0.18</td>
</tr>
<tr>
<td>Lag 7</td>
<td>0.12</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>Lag 8</td>
<td>0.12</td>
<td>0.19</td>
<td>0.16</td>
</tr>
<tr>
<td>Lag 9</td>
<td>0.12</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td>Lag 10</td>
<td>0.10</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>Lag 11</td>
<td>0.09</td>
<td>0.17</td>
<td>0.14</td>
</tr>
<tr>
<td>Lag 12</td>
<td>0.09</td>
<td>0.17</td>
<td>0.13</td>
</tr>
<tr>
<td>Lag 13</td>
<td>0.08</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td>Lag 14</td>
<td>0.08</td>
<td>0.14</td>
<td>0.12</td>
</tr>
<tr>
<td>Lag 15</td>
<td>0.07</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>L.B (15)</td>
<td>8943.34</td>
<td>9986.76</td>
<td>9910.23</td>
</tr>
</tbody>
</table>

6 Empirical Results

We start our analysis by estimating univariate Hawkes models for each maturity independently as shown in Table 4, where $\alpha$ is the impact factor of the self-excitation component, $\beta$ is the decay rate of the self-excitation effect, and $\mu$ is the baseline intensity which can be seen as the deterministic component of the non-homogenous Poisson process.\textsuperscript{14}

\textsuperscript{14}Duration between price changing trades are scaled by a factor of 60 in the construction of the point process for computational reasons.
Table 4: Univariate Hawkes estimation

<table>
<thead>
<tr>
<th>Variables</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.089</td>
<td>0.13</td>
<td>0.10</td>
</tr>
<tr>
<td>S.E.</td>
<td>(1.20E-5)</td>
<td>(3.24E-6)</td>
<td>(5.45E-6)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.129</td>
<td>0.154</td>
<td>0.12</td>
</tr>
<tr>
<td>S.E.</td>
<td>(2.00E-5)</td>
<td>(4.19E-6)</td>
<td>(7.28E-6)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.066</td>
<td>0.134</td>
<td>0.104</td>
</tr>
<tr>
<td>S.E.</td>
<td>(8.99E-6)</td>
<td>(6.15E-6)</td>
<td>(1.04E-5)</td>
</tr>
</tbody>
</table>

Panel B. Diagnostics

<table>
<thead>
<tr>
<th></th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>LogLikelihood</td>
<td>-72408.80</td>
<td>-132731.70</td>
<td>-121999.30</td>
</tr>
<tr>
<td>Observations</td>
<td>28644.00</td>
<td>116163.00</td>
<td>84651.00</td>
</tr>
<tr>
<td>Mean residual duration</td>
<td>0.9953</td>
<td>1.00</td>
<td>0.997</td>
</tr>
<tr>
<td>S.D. residual duration</td>
<td>0.96</td>
<td>0.976</td>
<td>0.956</td>
</tr>
<tr>
<td>L.B. test (20 lags)</td>
<td>284.60</td>
<td>106.45</td>
<td>56.14</td>
</tr>
</tbody>
</table>

All three securities exhibit strong autodependence as shown by the estimated $\alpha$’s with intensity half-lives of 5.12, 4.48, and 5.40 minutes for the 2-year, 5-year, and 30-year treasury securities respectively, which confirms the preliminary indication of strong autocorrelation found in Table 3 above. The residuals for all three series have a mean of 1 and a standard deviation that is close to 1, signalling that all three residuals are close to Exponentially (1) distributed. However, the test for independence using Ljung-Box test rejects the null-hypothesis that these residuals are independently distributed suggesting the need for a multi-variate model.

The multivariate Hawkes model is estimated by taking into account the interaction between trading activities at each maturity. Doing so allows us to investigate a much richer picture behind the forces driving the different parts of the yield curve. The model allows us to study the codependence or cross excitation of trading activities between treasury securities across time along with the self-excitation that will be reflected to some degree in the autocorrelation patterns reported above. Table 5 shows the estimated parameters for the multivariate Hawkes model as well as the satisfaction of the stationarity condition.

The columns denote the effect each column security has on the row security and the diagonal shows the self-excitation components. For example, the cross-excitation effect of a 5-year note price changing trade on a 2-year note price intensity is 0.001746. This suggests that the conditional probability of the trade occurrence in a 2-year note increases by 0.001746 on top of the baseline 0.0602 and self-excitation 0.08939 effects. This increase in the probability of occurrence of a price changing trade is, however, decaying at the speed of 0.6 units per minute.

The estimated parameters from the multivariate model show self-excitation effects that are close to those found in the univariate Hawkes model, but the standard errors of the estimated parameters suggest a significant influence of cross-trade effects in this multivariate framework. However, it is clear that the majority of price intensity is self-excited for all maturities. For instance, the cross-excitation terms in the 2-year note suggest that price movements are driven by trades occurring

---

15The estimation routine was written in C++ using OPT++ as the optimizing package. The log likelihood function is maximized using quasi-newton non-linear optimization. The C++ algorithm written for the calculation of the log likelihood has been tested against a simulated multivariate Poisson process. Empirical tests of the consistency of estimates have been successfully carried out. All estimated parameters attain convergence against the gradient tolerance of 1E-5 and the standard error is estimated using the inversion of the negative Hessian matrix. All the time of occurrence of price-changing trades is scaled in terms of minutes for numerical efficiency reasons. The multivariate Hawkes model is also expressed conditionally to improve on computational time. For more details, see Sohrmann and Tham (2006).
Table 5: Multivariate Hawkes estimation

<table>
<thead>
<tr>
<th>Variables</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A. 2-year bond</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.08939</td>
<td>0.001746</td>
<td>0.003673</td>
</tr>
<tr>
<td>S.E.</td>
<td>(1.25E-05)</td>
<td>(3.81E-06)</td>
<td>(5.06E-06)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1296</td>
<td>0.6</td>
<td>0.6397</td>
</tr>
<tr>
<td>S.E.</td>
<td>(2.10E-05)</td>
<td>(4.25E-06)</td>
<td>(4.29E-06)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.0602</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.E.</td>
<td>(1.02E-05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stationarity Condition $\frac{\alpha}{\beta} &lt; 1$</td>
<td>0.69</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>Panel B. 5-year bond</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.0003386</td>
<td>0.1309</td>
<td>0.0003843</td>
</tr>
<tr>
<td>S.E.</td>
<td>(1.08E-05)</td>
<td>(9.47E-06)</td>
<td>(9.61E-06)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.639</td>
<td>0.1549</td>
<td>0.5800</td>
</tr>
<tr>
<td>S.E.</td>
<td>(4.21E-06)</td>
<td>(1.23E-05)</td>
<td>(4.22E-06)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.1341</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.E.</td>
<td>(1.80E-05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stationarity Condition $\frac{\alpha}{\beta} &lt; 1$</td>
<td>0.003</td>
<td>0.22</td>
<td>0.0006</td>
</tr>
<tr>
<td>Panel C. 30-year note</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.007165</td>
<td>0.0007342</td>
<td>0.1003</td>
</tr>
<tr>
<td>S.E.</td>
<td>(7.48E-06)</td>
<td>(6.29E-06)</td>
<td>(7.77E-06)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.6392</td>
<td>0.6200</td>
<td>0.1201</td>
</tr>
<tr>
<td>S.E.</td>
<td>(4.26E-06)</td>
<td>(4.21E-06)</td>
<td>(9.79E-06)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.1007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.E.</td>
<td>(1.67E-05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stationarity Condition $\frac{\alpha}{\beta} &lt; 1$</td>
<td>0.011</td>
<td>0.0012</td>
<td>0.84</td>
</tr>
</tbody>
</table>
in the 5-year and 30-year markets but these effects are relatively much shorter lived and the cross impact effects much smaller. In each case, the decay pattern is slower for the own-market event compared to the cross-market effects.

However we can also see that each market is different as the price intensity is primarily driven by self-excitation effect and the coefficients of all impact and decay rates are significantly different from each other. This does suggest that the markets are possibly segmented to a degree and hence responding to different information, or there is more than one source of information that is driving the price process for each security. From the time deformation point of view, this also indicates that there are different time scales in each market and that the multivariate Hawkes model serves to coordinate activity between the different maturities. The estimated parameters also highlight that the long end of the yield curve has a significant impact on the shorter end and vice-versa. One can also observe that the all-weather 5-year note is a security that is the least affected by movements in other parts of the yield curve.

In brief, the markets do appear to be segmented to a considerable degree but not totally and there seem to be different ‘operational clocks’ or time scales at work in the treasury market. We will systematically test for the presence of a single market time scale using the likelihood ratio test and carry out various specification tests in the next section.

6.1 Diagnostic Tests and Tests for Time Deformation

As a first diagnostic, we examine if $\int_0^t \lambda^c (t; \mathcal{F}_t^2) = t$, which is effectively the null that there is time deformation. The most intuitive way of examining this is simply to plot $\int_0^t \lambda^c (t; \mathcal{F}_t^2)$ against clock time and see if it is linear and increases proportionately. Figures (6), (7), (8), (9), (10) and (11), show integrated price intensities and prices for each maturity across time on 30th June 2005; a day with an FOMC meeting. There was an increase of the base interest rate by 25 basis points. The median market expectation of the base rate rise is 25 basis points according to market survey from Informa Global Markets’ monthly report and it is interesting to observe prices for treasury notes and bonds decrease immediately and then rebound to a higher level after the announcement of the base rate rise. Such variation might be caused by the heterogeneous opinions about the scheduled interest rate announcement and the process of price discovery after an announcement. The presence of price discovery in the treasury market is also documented in Green (2004), where the uninformed participants in the market update their conditional expectation and learn about the full information price through the observed macroeconomic shock and order flow in the inter-dealer market. This price discovery process is the indirect flow of information into the financial market and is reflected in the price durations of treasury securities.

There is clearly a much greater time deformation impact on the two year note itself as the steepness of the time deformation function increases after the base rate announcement. The sudden increase in the slope of the integrated intensity, a measurement of trading activity, from 14:00 to 16:00 also coincides with the period of high volatility. These figures strengthen our view that integrated price intensity is a suitable operational clock.

In order to test the time deformation hypothesis more formally, we also carry out explicit parameter restriction tests on the estimated multivariate Hawkes model. Using Likelihood ratio (LR) tests we can test the following four hypotheses:

1. All three maturities have the same impact effects through the restriction that $H_0 : \alpha_{11} = \alpha_{22} = \alpha_{33}, \alpha_{12} = \alpha_{13} = \alpha_{21} = \alpha_{23} = \alpha_{31} = \alpha_{32}$

2. All three maturities have the same decay characteristics through the restriction that: $H_0 : \beta_{11} = \beta_{22} = \beta_{33}, \beta_{12} = \beta_{13} = \beta_{21} = \beta_{23} = \beta_{31} = \beta_{32}$

3. That each process a homogeneous Poisson Process: $H_0 : \alpha_{11} = \alpha_{22} = \alpha_{33} = \alpha_{12} = \alpha_{13} = \alpha_{21} = \alpha_{23} = \alpha_{31} = \alpha_{32} = 0$
Figure 6: Price of 2-year note vs clock time on 30th June 2005

Figure 7: Integrated Intensity of 2-year Notes vs time on 30th June 2005
Figure 8: Price of 5-year Notes vs time on 30th June 2005

Figure 9: Integrated Intensity of 5-year Notes Vs time on 30th June 2005
Figure 10: Price of 30-year Bond Vs time on 30th June 2005

Figure 11: Integrated Intensity of 30-year Bond Vs time on 30th June 2005
4. That the Multivariate Hawkes model does not dominate the univariate Hawkes models: $H_0$:
\[ \alpha_{12} = \alpha_{13} = \alpha_{21} = \alpha_{23} = \alpha_{31} = \alpha_{32} = 0 \]

Table 6 reports the relevant Likelihood Ratio Statistics for each hypothesis.

Table 6: Diagnostic and Likelihood ratio tests

<table>
<thead>
<tr>
<th>Variables</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Diagnostics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>28644</td>
<td>116163</td>
<td>84651</td>
</tr>
<tr>
<td>Mean residual duration</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>S.D. residuals duration</td>
<td>0.97</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>L.B. test (20 lags)</td>
<td>201.06</td>
<td>78.55</td>
<td>28.87</td>
</tr>
<tr>
<td><strong>Panel B. Likelihood Ratio Tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Log Likelihood of Multivariate Hawkes</td>
<td>-282717</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0 : \alpha_{11} = \alpha_{22} = \alpha_{33}.$</td>
<td>6868</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{12} = \alpha_{13} = \alpha_{21} = \alpha_{23} = \alpha_{31} = \alpha_{32}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0 : \beta_{11} = \beta_{22} = \beta_{33},$</td>
<td>112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{12} = \beta_{13} = \beta_{21} = \beta_{23} = \beta_{31} = \beta_{32}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0 : \text{Joint Homogeneous Poisson Process}$</td>
<td>92728</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0 : \text{Independent Univariate Hawkes Model}$</td>
<td>22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All the Likelihood Ratio statistics shown on the right hand side of Panel B reject the associated null hypothesis, in favor of our time deformation model. Panel B shows that the null of the same impact effects for each maturity is strongly rejected. Similarly it also indicates that the null hypothesis of the decay rates being the same is rejected. The null of the price duration being a homogenous Poisson process is also strongly rejected with a Likelihood Ratio of 92728 indicating the importance of the clustering effects of trade arrival and information flow. Finally when the restrictions are imposed to allow the multivariate Hawkes model to become three independent univariate Hawkes models at each maturity, the null is again rejected indicating the importance of the cross market effects but only with a much weaker significance level on this occasion. However, Panel A shows that we have residual autocorrelation indicating that a richer dynamic model is called for than we have been able to capture in our Multivariate Hawkes model.

This impression is supported with the results presented in Table 7, where we report formal tests of the model specification by applying the Engle Russell over-dispersion test for the generalized residuals which has a limiting Normal distribution along with a simple $\chi^2$ test for mean and standard deviation equivalence. The 5% critical value for the Engle Russell test is 1.645 and the null of an $Exp(1)$ distribution appears to be rejected for all maturities for both the univariate and multivariate generalized residual durations. The simple $\chi^2$ test indicates acceptance however and so we report in Panel C a specific Information Matrix based test for the $Exp(1)$ distribution given by Acosta and Rojas (2007) which clearly indicates acceptance of the Null of correct specification for both univariate and multivariate Hawkes models.

In the next section we turn to consider two robustness checks of the time deformation model by comparing the structural intensity-based volatility with the model-free realized volatility and then examining the distribution of bond returns standardized by the estimated instantaneous volatility.

### 6.2 Robustness Checks

#### 6.2.1 Realized vs. Intensity-based Volatility

As there has been very little work using this structural or “temporal” approach to measuring volatility using intensity-based point processes, we carry out a robustness check by studying the relationship
Table 7: Overdispersion tests

<table>
<thead>
<tr>
<th>Maturity</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Engle and Russell excess dispersion test</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate Hawkes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics</td>
<td>-2.37</td>
<td>-4.77</td>
<td>-8.06</td>
</tr>
<tr>
<td>Multivariate Hawkes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics</td>
<td>-3.54</td>
<td>-5.76</td>
<td>-8.72</td>
</tr>
<tr>
<td><strong>Panel B. $\chi^2$ test</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.94</td>
<td>0.95</td>
<td>0.92</td>
</tr>
<tr>
<td>Statistics</td>
<td>27033.32</td>
<td>110612.80</td>
<td>77694.36</td>
</tr>
<tr>
<td>P-value</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Panel C. Information Matrix test</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate Hawkes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics</td>
<td>3938.81</td>
<td>14164.28</td>
<td>10726.34</td>
</tr>
<tr>
<td>P-value</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Multivariate Hawkes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics</td>
<td>3951.78</td>
<td>14167.1</td>
<td>10715.41</td>
</tr>
<tr>
<td>P-value</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

between our measure and the now popular realized measure of volatility. The daily realized volatility is computed as the daily sum of returns squared with a sampling frequency of 30 minutes. The intensity-based volatility is calculated using equations (6)-(8) above so that the daily intensity based volatility is the integrated instantaneous volatility across everyday, using the integrated intensity multiplied by the price changes at different points in time. Figures (12), (13) and (14) show the high codependence between the two volatilities supporting the hypothesis of using trade-intensity as a directing process.

Table 8: Realized Vs Intensity-based Volatility

<table>
<thead>
<tr>
<th></th>
<th>2 year</th>
<th>5 year</th>
<th>30 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.8569</td>
<td>0.7954</td>
<td>0.4943</td>
</tr>
</tbody>
</table>

Table 8 indicates a relatively high degree of linear dependence between the two series at least for the 2 year and 5 year bonds. The realized volatility estimates are based on 20 observations aggregated every 30 minutes daily, whereas the intensity based measure is calculated using the integrated intensity across each day over the sample period.

The difference between the structural Hawkes based and the model-free realized volatility estimates are interesting and potentially important. Our results suggest a strong relation between these temporal and realized measures of volatility. The results support those findings regarding the impact of trade variables on variance using variance decomposition methods in the market microstructure literature. The relatively lower correlation coefficient of the 30-year bond could be related to the sharing of this market share between the electronic platform of eSpeed and BrokerTec. The differences between the intensity-based volatility and realized volatility are likely to be due to the market share of BrokerTec in the electronic inter-dealer market given that our measurement of volatility is based on the intensity of price changing trades in a section of the whole market. Pooling data from both eSpeed and BrokerTec might lead to a higher measure of dependence between the two measures. Alternatively, the results from our robustness tests reflect on the differences in the importance of

---

Figure 12: Realized Volatility Vs Intensity-Based Volatility - 2-year Note

Figure 13: Realized Volatility Vs Intensity-Based Volatility - 5-year Note
price discovery processes across maturities in the US treasury market.

6.2.2 Gaussianity tests

The second robustness test stems from Clark (1973), where we examine the adequacy of our time deformation model by conditioning the observed bond returns on the estimated instantaneous volatility from the Hawkes model. We do this by standardizing the bond returns with the estimated integrated instantaneous variance. Standardizing the observed leptokurtic return distribution of the observed returns with our operation clock should yield a Normal distribution for the standardized return series if a correct measure of information flow has been found and used as the operational clock in the time deformation model.

Table 9 provides in Panel A the $p$ values of various Normality tests on the non-conditioned bond returns and we can observe clear rejection of Gaussianity with strong Leptokurtosis particularly in the 2 year and 5 year bonds. Panel B reports the same test statistics for the conditioned bond returns and consistent with the theoretical argument our use of the intensity-based measure of volatility has returned a conditional Gaussian distribution in each case although there is relatively little effect on the 30 year returns.

Results from the robustness test supports the use of this temporal measure of volatility as the directing process of returns in the US treasury market. It also suggest that private information flows into the market through trades and is reflected in the arrival rates of price changing trades.
Table 9: Normality test for daily returns and time-transformed returns

<table>
<thead>
<tr>
<th>Panel A. Daily Log-Returns</th>
<th>Variables</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kurtosis</td>
<td>13.19</td>
<td>13.4</td>
<td><strong>2.45</strong></td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>1.99</td>
<td>1.71</td>
<td><strong>-0.28</strong></td>
</tr>
<tr>
<td></td>
<td>Shapiro-Wilk</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Kolmogorov-Smirnov</td>
<td>0.01</td>
<td>0.01</td>
<td><strong>0.15</strong></td>
</tr>
<tr>
<td></td>
<td>Cramer-von Mises</td>
<td>0.01</td>
<td>0.01</td>
<td><strong>0.07</strong></td>
</tr>
<tr>
<td></td>
<td>Anderson-Darling</td>
<td>0.01</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Bootstrapped Anderson-Darling</td>
<td>0.00</td>
<td>0.00</td>
<td><strong>0.078</strong></td>
</tr>
<tr>
<td></td>
<td>Bootstrapped Jarque-Bera</td>
<td>0.00</td>
<td>0.00</td>
<td><strong>0.28</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Daily Transformed Returns</th>
<th>Variables</th>
<th>2-year</th>
<th>5-year</th>
<th>30-year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kurtosis</td>
<td><strong>2.77</strong></td>
<td><strong>2.47</strong></td>
<td><strong>2.23</strong></td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-0.02</td>
<td>0.03</td>
<td><strong>-0.10</strong></td>
</tr>
<tr>
<td></td>
<td>Shapiro-Wilk</td>
<td>0.65</td>
<td>0.37</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Kolmogorov-Smirnov</td>
<td>0.15</td>
<td>0.15</td>
<td><strong>0.13</strong></td>
</tr>
<tr>
<td></td>
<td>Cramer-von Mises</td>
<td>0.21</td>
<td>0.25</td>
<td><strong>0.07</strong></td>
</tr>
<tr>
<td></td>
<td>Anderson-Darling</td>
<td>0.25</td>
<td>0.25</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Bootstrapped Anderson-Darling</td>
<td>0.14</td>
<td>0.18</td>
<td><strong>0.079</strong></td>
</tr>
<tr>
<td></td>
<td>Bootstrapped Jarque-Bera</td>
<td>0.46</td>
<td>0.385</td>
<td><strong>0.27</strong></td>
</tr>
</tbody>
</table>
7 Implications and Applications of Intensity-Based Time Deformation Model

Finding that treasury securities across different maturities operate on different ‘market clocks’ that have some degree of codependence is congruent with existing literature on term structure theory. Researchers in finance have studied the yield curve statistically and have found that shifts or changes in the shape of the yield curve are attributable to a few unobservable factors (Dai and Singleton, 2000). These affine macro-finance models, normally defined in no-arbitrage representations, provide macroeconomic interpretations for the latent variables. Our findings are consistent with this stream of the term structure literature in that these different "market clocks" are related to the different latent factors driving the yield curve but the yield curve itself is held together by the no-arbitrage restriction, which is effectively captured by the relationship between the self and cross excitation effects in our multivariate Hawkes model.

Another possible link to existing term structure theory is along the lines of the market segmentation (Culbertson, 1957) and preferred habitat (Modigliani and Sutch, 1966) models which rest on the view that investors differ in the maturities in which they are active given their liability structure. So for instance, Pension Funds may be more interested in long term bonds, say 30 years, while Life Assurance companies will be active in shorter maturities, perhaps 15 years and asset managers will be interested in shorter bonds. In both these theories, investors at each maturity are different and concentrate their activity on the specific maturity in which they are primarily concerned and so the bond price at each point in the yield curve will reflect the different liquidity structures and the separate forces of demand and supply at each maturity. The shape and movement of the yield curve is then seen to be determined by the different market forces and pressures at each maturity. Modern theories of the Yield Curve downplay the heterogeneity implied by these models and the different preference and information structure that is implied at each point in the yield curve. Vayanos and Vila (2007) have recently developed a modern Preferred Habitat model in which investors with preferences for specific maturities trade with risk averse arbitrageurs and it is the arbitrageurs who integrate the markets for different maturities by incorporating information about expected short rates into bond prices.

The relevance of integrated price intensity as a directing process for the return series supports the presence of information asymmetry and heterogeneous agents as well as the indirect information flow through trades which are seen as being important in market microstructure theory. Information asymmetry and heterogeneous agents induce a gradual price discovery process that has a non-homogeneous arrival rate of trades as private information is incorporated into the market. In particular, the price impact of a trade tends to increase as the time duration between two price changing trades decreases, suggesting that an increased trading activity would be associated with a higher level of information asymmetry. These effects allow one to hypothesize that in the event of uncertainty and the presence of asymmetric information, price intensity in the market would increase.

While our model provides an alternative study to price discovery and term structure theory in the treasury market, it can also be used as an alternative pricing model for interest rate related derivatives as shown in the appendix. Such an exercise provides a “temporal” derivatives pricing model with a market microstructure flavor as an alternative to the existing pricing models.

8 Conclusion

In this paper, we have addressed the issue of time deformation in the US Treasury market and propose a new continuous time multi-variate time deformation model motivated by market microstructure theory. Our objective has been to study the interrelationship between the different asset price dynamics from the perspective of the frequency of price changing trades. In this way we have been able to determine the stochastic nature of information arrival in different maturities of the yield curve.
and hence model the price discovery process using a multi-variate point process. In doing this we have been able to uncover the underlying market time structure within the yield curve and recover normality for the asset returns using our estimated directing process; a result that is consistent with market microstructure theory.

We find that integrated price intensity is an appropriate directing process in US Treasury market suggesting that private information is revealed indirectly through trades in the presence of information asymmetry and heterogeneous agents. Our results also show that US treasury notes and bonds are driven by different ‘market clocks’, but are related to each other through the cross-excitation effects of trades at different maturities. This finding is consistent with macro finance models which suggest the presence of multiple factors driving the yield curve under a no-arbitrage restriction. Our result is also consistent with modern preferred habitat theory of the yield curve suggesting the presence of different classes of participants trading with risk averse arbitrageurs across securities of different maturities. The model specification and econometric techniques used would seem to open up alternative channels for understanding and modeling term structure theory as well as the pricing of interest rate related derivatives.
References


[66] Vayanos D. and Vila Jean-Luc, A preferred habitat model of the term structure of interest rates, mimeo LSE.
9 Appendix

This section demonstrates the application of our model in derivative pricing. We could model bond prices using the HJM stochastic differential equation as described below with the instantaneous volatilities determined by the Hawkes models given above. Suppose that the forward rate is given by \( df_t(T) = \alpha(t, T)dt + \sum_{i=1}^{n} \sigma_i(t, T)dz_{i,t}, \) as usual, where \( \sigma_i(t, T), \ i = 1, \ldots, n, \) are Gaussian. Set \( a_i(t, T) = -\int_t^T \sigma_i(t, s)ds, \ i = 1, \ldots, n. \) Pure discount bond prices follow the process

\[
\frac{dB_i(T)}{B_i(T)} = (r_t + b(t, T)) dt + a(t, T) dz_t, \tag{22}
\]

under the objective measure \( Q, \) where

\[
a = (a_1, \ldots, a_n)', \tag{23}
\]

\[
a_i(t, T, \omega) = -\int_t^T \sigma_i(t, s, \omega)ds, \ i = 1, \ldots, n, \tag{24}
\]

\[
b(t, T, \omega) = -\int_t^T \alpha(t, s, \omega)ds + \frac{1}{2} \sum_{i=1}^{n} a_i^2(t, T, \omega). \tag{25}
\]

When forward rate volatilities are Gaussian it possible to obtain formulae for some simple instruments. Brenner and Jarrow (93) [?] and Au and Thurston (94) [?] showed that there is a Black’s formula for \( c_T(T_1, T_2), \)

\[
a_T(T_1, T_2) = B_T(T_2)N(d) - XB_T(T_1)N(d - w) \tag{26}
\]

where

\[
d = \frac{1}{\sqrt{w}} \ln \left( \frac{B_T(T_2)}{XB_T(T_1)} \right) + \frac{1}{2} \sqrt{w}, \tag{27}
\]

\[
w = \sum_{i=1}^{n} \int_t^{T_1} (a_i(u, T_2) - a_i(u, T_1))^2 du, \tag{28}
\]

and the initial forward curve has been fitted to match the market values \( B_t(T) \) of PDBs.

Suppose that the bond volatility curve is now fitted by a curve of Nelson and Siegel type so that

\[
a(t) = \beta_0 + (\beta_1 + \beta_2 \tau)e^{-k\tau} \tag{29}
\]

where \( \tau = T - u \) and \( a(t) \) is again given by the estimated intensity model (7) using the Hawkes process above. Then

\[
w = \int_t^{T_1} (a_i(u, T_2) - a_i(u, T_1))^2 du \tag{30}
\]

\[
= \int_t^{T_1} \left( \beta_0 + (\beta_1 + \beta_2 (T_2 - u))e^{-k(T_2-u)} - \beta_0 - (\beta_1 + \beta_2 (T_1 - u)) e^{-k(T_1-u)} \right)^2 du \tag{31}
\]

\[
= \int_t^{T_1} \left( \frac{(\beta_1 + \beta_2 (T_2 - u))^2 e^{-2k(T_2-u)}}{2} + (\beta_1 + \beta_2 (T_1 - u))^2 e^{-2k(T_1-u)} - 2(\beta_1 + \beta_2 (T_2 - u))(\beta_1 + \beta_2 (T_1 - u)) e^{-k(T_1+T_2-2u)} \right) du \tag{32}
\]

\[
= \int_t^{T_1} (\beta_1 + \beta_2 (T_2 - u))^2 e^{-2k(T_2-u)} du \tag{33}
\]

\[
+ \int_t^{T_1} (\beta_1 + \beta_2 (T_1 - u))^2 e^{-2k(T_1-u)} du \tag{34}
\]

\[
-2 \int_t^{T_1} (\beta_1 + \beta_2 (T_2 - u))(\beta_1 + \beta_2 (T_1 - u)) e^{-k(T_1+T_2-2u)} du \tag{35}
\]
But

\[
\int u e^{au} du = e^{au} \left( \frac{u}{a} - \frac{1}{a^2} \right)
\]

(36)

\[
\int u^2 e^{au} du = e^{au} \left( \frac{u^2}{a} - \frac{2u}{a^2} + \frac{1}{a^3} \right)
\]

(37)

so

\[
\int (\beta_1 + \beta_2 (T_2 - u))^2 e^{-2k(T_2 - u)} du
\]

(38)

\[
e^{-2kT_2} \int_1^{T_1} (\beta_1 + \beta_2 T_2)^2 + \beta_2^2 u^2 - 2 (\beta_1 + \beta_2 T_2) \beta_2 u \right) e^{2ku} du
\]

(39)

\[
e^{-2kT_2} (\beta_1 + \beta_2 T_2)^2 \frac{e^{2ku}}{2k}
\]

(40)

\[
+e^{-2kT_2} \beta_2 e^{2ku} \left( \frac{u^2}{2k} - \frac{2u}{(2k)^2} + \frac{2}{(2k)^3} \right)
\]

(41)

\[-2e^{-2kT_2} (\beta_1 + \beta_2 T_2) \beta_2 e^{2ku} \left( \frac{u}{2k} - \frac{1}{(2k)^2} \right)
\]

(42)

and

\[
\int (\beta_1 + \beta_2 (T_2 - u))^2 e^{-2k(T_2 - u)} du
\]

(43)

\[
e^{-2kT_1} (\beta_1 + \beta_2 T_1)^2 \frac{e^{2ku}}{2k}
\]

(44)

\[
+\beta_2^2 e^{-2kT_1} e^{2ku} \left( \frac{u^2}{2k} - \frac{2u}{(2k)^2} + \frac{2}{(2k)^3} \right)
\]

(45)

\[-2e^{-2kT_1} (\beta_1 + \beta_2 T_1) \beta_2 e^{2ku} \left( \frac{u}{2k} - \frac{1}{(2k)^2} \right)
\]

(46)

and

\[
\int ((\beta_1 + \beta_2 T_2) - \beta_2 u) ((\beta_1 + \beta_2 T_1) - \beta_2 u) e^{-k(T_1 + T_2 - 2u)} du
\]

(47)

\[
e^{-k(T_1+T_2)} (\beta_1 + \beta_2 T_2) (\beta_1 + \beta_2 T_1) \int e^{2ku} du
\]

(48)

\[
+e^{-k(T_1+T_2)} \beta_2 \int u^2 e^{2ku} du
\]

(49)

\[-e^{-k(T_1+T_2)} (2\beta_1 + \beta_2 (T_1 + T_2)) \beta_2 \int u e^{2ku} du
\]

(50)

\[
e^{-k(T_1+T_2)} (\beta_1 + \beta_2 T_2) (\beta_1 + \beta_2 T_1) \frac{e^{2ku}}{2k}
\]

(51)

\[
+e^{-k(T_1+T_2)} \beta_2^2 e^{2ku} \left( \frac{u^2}{2k} - \frac{2u}{(2k)^2} + \frac{2}{(2k)^3} \right)
\]

(52)

\[-e^{-k(T_1+T_2)} (2\beta_1 + \beta_2 (T_1 + T_2)) \beta_2 e^{2ku} \left( \frac{u}{2k} - \frac{1}{(2k)^2} \right)
\]

(53)

33
So

\[ w = \begin{bmatrix}
\sum e^{-2kT_2} (\beta_1 + \beta_2 T_2)^2 \frac{e^{2ku}}{2k} \\
+ e^{-2kT_2} \beta_2 e^{2ku} \left( \frac{u^2}{2k} - \frac{2u}{(2k)^2} + \frac{2}{(2k)^4} \right) \\
-2e^{-2kT_2} (\beta_1 + \beta_2 T_2) \beta_2 e^{2ku} \left( \frac{u}{2k} - \frac{1}{(2k)^2} \right) \\
+ e^{-2kT_1} (\beta_1 + \beta_2 T_1)^2 \frac{e^{2ku}}{2k} \\
+ \beta_2^2 e^{-2kT_1} e^{2ku} \left( \frac{u^2}{2k} - \frac{2u}{(2k)^2} + \frac{2}{(2k)^4} \right) \\
-2e^{-2kT_1} (\beta_1 + \beta_2 T_1) \beta_2 e^{2ku} \left( \frac{u}{2k} - \frac{1}{(2k)^2} \right) \\
-2e^{-k(T_1+T_2)} (\beta_1 + \beta_2 T_2) (\beta_1 + \beta_2 T_2) \frac{e^{2ku}}{2k} \\
-2e^{-k(T_1+T_2)} \beta_2 e^{2ku} \left( \frac{u^2}{2k} - \frac{2u}{(2k)^2} + \frac{2}{(2k)^4} \right) \\
+ 2e^{-k(T_1+T_2)} (2\beta_1 + \beta_2 (T_1 + T_2)) \beta_2 e^{2ku} \left( \frac{u}{2k} - \frac{1}{(2k)^2} \right)
\end{bmatrix} \] 

(54)

A simple empirical exercise is demonstrated here. If we approximate the instantaneous volatility for all maturities and fit a Nelson-Siegel type function to these volatilities contemporaneously. We can use the estimated parameters from the Nelson-Siegel function to price caps which can be stripped into caplets and spot rates across all maturities can be backout from the caplets.