

Modifying Gaussian term structure models when interest rates are near the zero lower bound

Leo Krippner*

October 26, 2011

Abstract

With nominal interest rates currently at or near their zero lower bound (ZLB) in many major economies, it has become untenable to apply Gaussian affine term structure models (GATSMs) while ignoring their inherent non-zero probabilities of negative interest rates. In this article I modify GATSMs by representing physical currency as call options on bonds to establish the ZLB. The resulting ZLB-GATSM framework remains tractable, producing a simple closed-form analytic expression for forward rates and requiring only elementary univariate numerical integration (over time to maturity) to obtain interest rates and bond prices. I demonstrate the salient features of the ZLB-GATSM framework using a two-factor model. An illustrative application to U.S. term structure data indicates that movements in the model state variables have been consistent with unconventional monetary policy easings undertaken after the U.S. policy rate reached the ZLB in late 2008.

JEL: E43, G12, G13

Keywords: zero lower bound; term structure of interest rates; Gaussian affine term structure models.

1 Introduction

In this article I develop a generic framework for imposing the zero lower bound (ZLB) for nominal interest rates on Gaussian affine term structure models (GATSMs).

My primary motivation for developing the ZLB-GATSM framework is to allow GATSM dynamics to be used for modeling the term structure in a theoretically consistent manner in all interest rate environments, as I will discuss further below. At the same time, I seek to preserve as much as possible two key features that have made GATSMs extremely popular; i.e. their flexibility and tractability. Specifically, GATSMs may be specified arbitrarily (in terms of the number of factors and inter-factor relationships) while retaining closed-form analytic solutions for pricing standard interest rate instruments, and multivariate normal transition densities. Those features make GATSMs straightforward to estimate and apply relative to other term structure

*Reserve Bank of New Zealand and CAMA. Email: leo.krippner@rbnz.govt.nz. I thank Iris Claus, Francis Diebold, Pedro Gomis, Michelle Lewis, Glenn Rudebusch, Christie Smith, and participants at a Reserve Bank of New Zealand presentation for helpful comments.

models.¹ To briefly highlight the popularity of GATSMs, which includes arbitrage-free versions of the Nelson and Siegel (1987) class,² Hamilton and Wu (2010) introduces them as “the basic workhorse in macroeconomics and finance” and notes seven recent examples of their application. Rudebusch (2010) surveys GATSM applications in macrofinance.

However, it is well-acknowledged (e.g. see Piazzesi (2010) p. 716) that GATSMs cannot be theoretically consistent in the “real world” where physical currency is available. The inconsistency arises because any unconstrained Gaussian process for short rate dynamics implies non-zero probabilities of negative interest rates. On such a realization, one could realize an arbitrage profit by borrowing (therefore receiving the absolute interest rate) to buy and hold physical currency (with a known return of zero). Alternatively, one could sell bond options based on the non-zero probabilities of negative interest rates in GATSMs, but with no probability of an out-of-the-money expiry in practice.

Despite their theoretical inconsistency, GATSMs are often applied with the assumption (as discussed in Piazzesi (2010) p. 716, but often left implicit) that the inherent probabilities of negative interest rates in GATSMs are sufficiently small to make the model immaterially different to a “real world” model subject to the ZLB. While that has usually been the case over history, such an assumption is obviously untenable in several major developed economies at the time of writing. Near-zero policy interest rates have been maintained in the United States and the United Kingdom since late-2008/early-2009, and in Japan since the 1990s. Each of those countries have also engaged in unconventional monetary policy easing due to the ZLB constraint.

The material probability of negative interest rates in GATSMs in near-zero interest rate environments amounts to model mis-specification. In essence, if term structure data are materially constrained by the ZLB but the GATSM applied to the data assumes no constraints, then the estimated GATSM and its state variables cannot provide a valid representation of the term structure and its dynamics. The mis-specification will affect even routine GATSM applications, such as monitoring the level and shape of the estimated term structure to provide a gauge of the stance of monetary policy. The implications are compounded for any relationships established via GATSMs between term structure and macroeconomic data (e.g. measures of inflation and real output growth), because macroeconomic data are not constrained to be non-negative.

There is a prevalent literature on term structure models with dynamics designed to avoid negative interest rates, but these inevitably come at the cost of flexibility and/or tractability relative to GATSMs. For example, closed-form analytic solutions and transition densities are not available for arbitrary specifications of multifactor Cox, Ingersoll, and Ross (1985b)/square-root models, and even the special case of independent factors requires the product of noncentral chi-square distributions.³

A more general framework that avoids negative interest rates is due to Black (1995), who makes the observation that physical currency effectively provides an option $\underline{r}(t) =$

¹GATSMs have in the past been subject to estimation issues, unrelated to their Gaussian dynamics, but Hamilton and Wu (2010) and Joslin, Singleton, and Zhu (2010) offer reliable estimation methods.

²Christensen, Diebold, and Rudebusch (2009, 2011) derive their arbitrage-free Nelson and Siegel (1987) models via GATSM specifications that reproduce Nelson and Siegel (1987) factor loadings.

³See Piazzesi (2010) p. 727. James and Webber (2000) pp. 226-233 discusses a variety of positive interest rate models.

$\max[r(t), 0]$, where $\underline{r}(t)$ is the “real world” short rate and $r(t)$ is the “shadow short rate”; i.e. the short rate that would prevail, including negative values, if physical currency were not available.

Several authors have applied the direct Black (1995) approach to specific GATSMs, but the resulting models have limited flexibility and/or tractability. For example, Gorovoi and Linetsky (2004) obtains an analytic expression for the ZLB-Vasicek (1977) model, but the approach does not appear to generalize to multiple factors.⁴ Ueno, Baba, and Sakurai (2006), Ichiue and Ueno (2007), and Kim and Singleton (2011) respectively use Monte Carlo simulations, interest rate lattices, and finite-difference grids to illustrate and/or estimate two-factor ZLB-GATSMs.⁵ However, the tractability of such numerical methods would decline with more factors.

The framework I develop in this article follows the concept of Black (1995), but is based on bond and bond option prices rather than directly on short rates. Essentially, if a GATSM is used to represent the shadow term structure, then the ZLB-GATSM term structure can be expressed in terms of shadow forward bonds and call options on those bonds to represent the optionality provided by physical currency. I present the set-up intuitively over finite steps of time to maturity and then take the infinitesimal limit to obtain continuous-time ZLB-GATSM forward rates. The latter turns out to be a very simple closed-form analytic expression, involving just exponential functions and the univariate cumulative normal distribution. ZLB-GATSM interest rates are calculated via numerical integration of the forward rate function, but I show that the integration will always be univariate (over time to maturity) regardless of the number of factors in the shadow GATSM.

I illustrate the salient features of the ZLB-GATSM framework using a calibrated two-factor GATSM for the shadow term structure. The first exercise shows how the ZLB-GATSM framework transforms negative GATSM/shadow forward and interest rates into non-negative forward and interest rates. The second exercise uses an illustrative application to U.S. term structure data to highlight how the model-implied shadow short rate moves to increasingly negative values around the announcements of unconventional monetary policy easings undertaken after the U.S. policy rate reached the ZLB in late 2008.

The outline of the remainder of the article is as follows. Section 2 outlines the intuition of my approach to developing a ZLB term structure framework. Section 3 provides the formal derivation for the generic ZLB-GATSM framework. Section 4 contains the illustrative exercises, and section 5 briefly discusses potential extensions of the ZLB-GATSM framework introduced in this article. Section 6 concludes.

2 A general ZLB framework

In this section I outline the intuition of my approach to developing a ZLB term structure framework. I present the concepts without reference to a particular model of the term structure and in a relatively informal manner, in particular leaving some aspects to

⁴See Kim and Singleton (2011) p. 11. Ichiue and Ueno (2006) and Ueno, Baba, and Sakurai (2006) apply the Gorovoi and Linetsky (2004) model to the Japanese government bond market.

⁵Kim and Singleton (2011) is actually a two-factor quadratic Gaussian model, but it embeds the two-factor GATSM as a special case.

formalize in section 3.

As a familiar starting point, section 2.1 details how forward rates may be determined from (zero-coupon) bonds priced with a term structure when there is no influence from physical currency; i.e. under the shadow term structure. The framework is initially developed using finite time/time-to-maturity steps, which I hereafter abbreviate to “finite-steps”. Calculating the infinitesimal limit for the finite-step expression gives the continuous-time expression for shadow forward rates. While the result might seem tautological, it makes the subsequent derivation of ZLB forward rates clearer.

Section 2.2 discusses how the availability of physical currency effectively offers investors, i.e. agents with surplus funds,⁶ a series of finite-step European call options on bonds (which I hereafter abbreviate to “options”). Section 2.3 shows how those options modify the finite-step shadow forward rates from section 2.1 to give finite-step ZLB forward rates. Calculating the infinitesimal limit for finite-step ZLB forward rates gives the continuous-time expression for ZLB forward rates.

2.1 The term structure without physical currency

2.1.1 Finite-step set-up

When no physical currency is available, investors can only invest their surplus funds in securities priced from the prevailing term structure, regardless of the prospective returns on offer. The fixed interest class of securities includes instruments such as “cash”/“liquidity” and “overnight/on-call deposits”, which are bonds with very short times to maturity. I introduce the notation $P(t, \delta)$ to represent such bonds, which settle for price $P(t, \delta)$ at time t and pay 1 at time $t + \delta$, where $\delta > 0$ is a finite-step representing the time to maturity. The annualized rate of return over the finite-step on a continuously-compounding basis is the finite-step short rate:

$$\begin{aligned} R(t, \delta) &= \frac{1}{\delta} \log \left[\frac{1}{P(t, \delta)} \right] \\ &= -\frac{1}{\delta} \log [P(t, \delta)] \end{aligned} \tag{1}$$

Bonds with longer maturities may be decomposed into an equivalent portfolio of a finite-step bond $P(t, \delta)$ and a series of forward contracts on finite-step bonds (hereafter abbreviated to “forward bonds”). I therefore use $P(t, \delta)$ and forward bonds as my basic building blocks for the term structure.

Table 1 considers the valuation, as at time t , of a forward bond under an arbitrary term structure. The cashflows are the forward bond price $P(t, \tau, \tau + \delta)$ due at the time of settlement $t + \tau$, and a return of 1 at the time of maturity $t + \tau + \delta$, where τ is a non-negative integer multiple of δ representing a future horizon from time t . The forward bond cashflows are non-contingent, and so the respective discount factors are simply the bond prices $P(t, \tau)$ and $P(t, \tau + \delta)$ from the prevailing term structure. Those bonds settle at time t and pay 1 at the respective times of maturity $t + \tau$ and $t + \tau + \delta$ (i.e. times to maturity $t + \tau$ and $t + \tau + \delta$), where $\tau \geq 0$.

⁶I typically take the perspective of the investor, but all comments and calculations from the investor’s perspective can be reversed to give the borrower’s perspective.

Table 1: Forward bond without physical currency		
Time	$t + \tau$	$t + \tau + \delta$
Cashflows	$-\text{P}(t, \tau, \tau + \delta)$	1
Discount factors	$\text{P}(t, \tau)$	$\text{P}(t, \tau + \delta)$
Present values	$-\text{P}(t, \tau) \cdot \text{P}(t, \tau, \tau + \delta)$	$\text{P}(t, \tau + \delta)$

The present value of the forward bond cashflows must equal zero to preclude arbitrage opportunities, and that results in an arbitrage-free expression for the forward bond in terms of bond prices. That is $-\text{P}(t, \tau) \cdot \text{P}(t, \tau, \tau + \delta) + \text{P}(t, \tau + \delta) = 0$, which leads to the standard expression:⁷

$$\text{P}(t, \tau, \tau + \delta) = \frac{\text{P}(t, \tau + \delta)}{\text{P}(t, \tau)} \quad (2)$$

The associated finite-step forward rate $F(t, \tau, \tau + \delta)$, i.e. the forward rate from time t to $t + \delta$ expressed on a continuously-compounding basis, is:

$$\begin{aligned} F(t, \tau, \tau + \delta) &= \frac{1}{\delta} \log \left[\frac{1}{\text{P}(t, \tau, \tau + \delta)} \right] \\ &= -\frac{1}{\delta} \log [\text{P}(t, \tau, \tau + \delta)] \\ &= -\frac{1}{\delta} \log \left[\frac{\text{P}(t, \tau + \delta)}{\text{P}(t, \tau)} \right] \end{aligned} \quad (3)$$

The finite-step short rate $R(t, \delta)$ is the finite-step forward rate at $\tau = 0$, i.e. $R(t, \delta) = F(t, 0, \delta)$.⁸

2.1.2 Continuous-time shadow forward rate

The instantaneous forward rate $f(t, \tau)$ may be calculated as the infinitesimal limit of finite-step forward rate $F(t, \tau, \tau + \delta)$; i.e.:

$$\begin{aligned} f(t, \tau) &= \lim_{\delta \rightarrow 0} F(t, \tau, \tau + \delta) \\ &= \lim_{\delta \rightarrow 0} \left\{ -\frac{1}{\delta} \log \left[\frac{\text{P}(t, \tau + \delta)}{\text{P}(t, \tau)} \right] \right\} \\ \langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ -\frac{d}{d\delta} \left(\log \left[\frac{\text{P}(t, \tau + \delta)}{\text{P}(t, \tau)} \right] \right) \right\}}{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \delta \right\}} \end{aligned} \quad (4)$$

⁷See, for example, Filipović (2009) p. 7. Note here and in subsequent footnotes that Filipović (2009) uses time t and time of maturity T notation, which is equivalent to my time t and time to maturity τ notation with $T = t + \tau$. James and Webber (2000) is another useful textbook reference for the standard term structure relationships I use in this article.

⁸Bonds could be equivalently written as forward bonds with zero time to settlement; i.e. $\text{P}(t, \tau) = \text{P}(t, 0, \tau)$ and $\text{P}(t, \tau + \delta) = \text{P}(t, 0, \tau + \delta)$.

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \delta \right\} = \lim_{\delta \rightarrow 0} \{1\} = 1$, and the derivative in the numerator of equation 4 is:

$$\begin{aligned}
\frac{d}{d\delta} \left(-\log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right) &= \frac{d}{d\delta} (-\log [P(t, \tau + \delta)] + \log [P(t, \tau)]) \\
&= -\frac{d}{d\delta} \log [P(t, \tau + \delta)] + 0 \\
\langle \text{Chain rule} \rangle &= -\frac{d}{d(\tau + \delta)} \log [P(t, \tau + \delta)] \frac{d}{d\delta} (\tau + \delta) \\
&= f(t, \tau + \delta)
\end{aligned} \tag{5}$$

where $f(t, \tau + \delta) = \frac{d}{d(\tau + \delta)} \log [P(t, \tau + \delta)]$ is the standard term structure definition relating bond prices to forward rates,⁹ and $\frac{d}{d\delta} (\tau + \delta) = 1$. The final result for equation 4 is therefore:

$$\begin{aligned}
f(t, \tau) &= \lim_{\delta \rightarrow 0} \{f(t, \tau + \delta)\} \\
&= f(t, \tau)
\end{aligned} \tag{6}$$

Note that the instantaneous short rate is $\lim_{\delta \rightarrow 0} R(t, \delta) = \lim_{\delta \rightarrow 0} F(t, 0, \delta) = f(t, 0) = r(t)$.

The instantaneous forward rate $f(t, \tau)$ may then be used to establish (continuously compounded zero-coupon) interest rates and bond price expressions in the usual way:¹⁰

$$r(t, \tau) = \frac{1}{\tau} \int_0^\tau f(t, v) dv \tag{7}$$

and:

$$P(t, \tau) = \exp[-\tau \cdot r(t, \tau)] \tag{8}$$

2.2 Physical currency and the shadow term structure

Physical currency may be viewed as a security that offers a zero return. So if the finite-step shadow bond $P(t, \delta)$ offers negative returns, investors can choose to hold their surplus funds as physical currency to at least obtain a zero return. That observation was originally made by Black (1995), who therefore proposed that physical currency offers an option on the shadow short rate itself, i.e.: $\underline{r}(t) = \max[r(t), 0]$, where $\underline{r}(t)$ is the “real world” short rate given the availability of physical currency, and $r(t)$ is the shadow short rate, i.e. the short rate when physical currency is not available.

For the ZLB framework in this article, I detail how physical currency relates to bonds and options priced using the shadow term structure. Of course, the short rate optionality and bond/option perspectives on physical currency should be equivalent, and I explicitly show that to be case in what follows.

⁹See, for example, Filipović (2009) p. 7.

¹⁰See, for example, Filipović (2009) p. 7. Filipović (2009) actually uses the equivalent expressions $r(t, T) = -\frac{1}{T-t} \log [P(t, T)]$ and $P(t, T) = \exp \left[-\int_t^T f(t, u) du \right]$, where $T = t + \tau$ is the time of maturity and $u = t + v$. Therefore $v = u - t$, $dv = du$, and the limits of integration with respect to v become $t - t = 0$ and $T - t = \tau$.

Investing in a finite-step shadow bond with a price $P(t, \delta) = 1$ is equivalent to holding a unit of physical currency over the finite-step, because the subsequent return of 1 in both cases results in a zero interest rate. Explicitly, if $P(t, \delta) = 1$, then the finite-step shadow short rate is:

$$\begin{aligned}
R(t, \delta) &= -\frac{1}{\delta} \log [P(t, \delta)] \\
&= -\frac{1}{\delta} \log [1] \\
&= -\frac{1}{\delta} \cdot 0 \\
&= 0
\end{aligned} \tag{9}$$

For the discussion below, it is also worthwhile bearing in mind the corollaries that if $P(t, \delta) > 1$ then $R(t, \delta) < 0$, and if $P(t, \delta) < 1$ then $R(t, \delta) > 0$.

The availability of physical currency is therefore equivalent to finite-step shadow bonds with a price $P(t, \delta) = 1$ being available as an investment, if desired. The “if desired” is the crucial turn of phrase that underlies the optionality of physical currency; i.e. investors always have the right but not the obligation to hold it. Of course, if the prevailing market price of shadow bonds $P(t, \delta)$ is greater than 1, then investors will choose to maximize their return by holding physical currency at $P(t, \delta) = 1$. Alternatively, if the prevailing market price of bonds $P(t, \delta)$ is less than 1, then investors will choose to hold the bonds rather than substituting physical currency at $P(t, \delta) = 1$.

In summary, the investor will choose $\underline{P}(t, \delta) = \min \{1, P(t, \delta)\}$, where $\underline{P}(t, \delta)$ is my notation for the finite-step ZLB bond. The optionality arising from the presence of physical currency establishes the lower bound of zero for the finite-step ZLB short rate $\underline{R}(t, \delta)$. Explicitly:

$$\begin{aligned}
\underline{R}(t, \delta) &= \frac{1}{\delta} \log \left[\frac{1}{\underline{P}(t, \delta)} \right] \\
&= -\frac{1}{\delta} \log [\underline{P}(t, \delta)] \\
&= -\frac{1}{\delta} \log [\min \{1, P(t, \delta)\}] \\
&= \left[\max \left\{ -\frac{1}{\delta} \log [1], -\frac{1}{\delta} \log [P(t, \delta)] \right\} \right] \\
&= \max \{0, R(t, \delta)\}
\end{aligned} \tag{10}$$

So $\underline{R}(t, \delta) = 0$ if $R(t, \delta) < 0$, otherwise $\underline{R}(t, \delta) = R(t, \delta)$. Taking the limit reproduces the Black (1995) concept of physical currency being an option on the short rate; i.e. $\lim_{\delta \rightarrow 0} \underline{R}(t, \delta) = \max \{\lim_{\delta \rightarrow 0} R(t, \delta), 0\}$, so $\underline{r}(t) = \max [r(t), 0]$.

$\underline{P}(t, \delta)$ may be more conveniently expressed as the sum of $P(t, \delta)$ and the payoff for an option on $P(t, \delta)$, i.e.:

$$\begin{aligned}
\underline{P}(t, \delta) &= \min \{1, P(t, \delta)\} \\
&= P(t, \delta) + \min \{1 - P(t, \delta), 0\} \\
&= P(t, \delta) - \max \{P(t, \delta) - 1, 0\} \\
&= P(t, \delta) - C(t, 0, \delta)
\end{aligned} \tag{11}$$

where $C(t, 0, \delta) = \max\{P(t, \delta) - 1, 0\}$ is the payoff for an option, with immediate expiry and a strike price of 1, on the finite-step bond $P(t, \delta)$.

Beyond the first finite-step, investors also know that physical currency will be available at any point of time in the future, and so will always offer them the choice between holding surplus funds as physical currency or investing the funds at the prevailing market price for shadow bonds. In notation, at a future time $t + \tau$ (where $\tau \geq 0$ is the horizon from time t), the finite-step ZLB bond is:

$$\begin{aligned} \underline{P}(t + \tau, \delta) &= \min\{1, P(t + \tau, \delta)\} \\ &= P(t + \tau, \delta) - C(t + \tau, 0, \tau + \delta) \end{aligned} \quad (12)$$

While the optionality inherent in $\underline{P}(t + \tau, \delta)$ is analogous to that for $\underline{P}(t, \delta)$, the difference is that $\underline{P}(t, \delta)$ is non-contingent (because the payoff from $C(t, 0, \delta)$ is already known at time t) while $\underline{P}(t + \tau, \delta)$ is contingent (because the payoff from $C(t + \tau, 0, \tau + \delta)$ is unknown at time t , being dependent on the price of $P(t + \tau, \delta)$ at the future time $t + \tau$). Nevertheless, the price of an option with a non-zero time to expiry τ may be obtained by calculating its expected value, as at time t , of the payoff at time $t + \tau$. Specifically, Filipović (2009) p. 109 gives the option price as:

$$C(t, \tau, \tau + \delta) = \mathbb{E}_t \left\{ \exp \left(- \int_0^\tau r(v) dv \right) \cdot \max [P(t + \tau, \delta) - 1, 0] \right\} \quad (13)$$

where $C(t, \tau, \tau + \delta)$ is the value, as at time t , of the option, expiring at time $t + \tau$ with a strike price of 1, on the bond $P(t + \tau, \delta)$.

Note that equation 13 is effectively discounting $\max [P(t + \tau, \delta) - 1, 0]$, the contingent cashflow, to its present value with the shadow short rate $r(t)$ rather than the ZLB short rate $\underline{r}(t)$. The distinction is important and it is used again in the following section, but I leave the formal justification to section 3.

2.3 The term structure with physical currency

Section 2.1 introduced the shadow forward rates $F(t, \tau, \tau + \delta)$ and $f(t, \tau)$ that would prevail in a term structure without physical currency, while section 2.2 showed that the current and future availability of physical currency may be represented as a series of options $C(t, \tau, \tau + \delta)$ available to investors. This section combines those two elements to give the term structure that should prevail when physical currency is available.

2.3.1 Finite-step set-up

The intuition for my ZLB framework can be expressed succinctly and intuitively as follows: an investor will only purchase the finite-step ZLB forward bond $\underline{P}(t, \tau, \tau + \delta)$ if the price results in compensation $C(t, \tau, \tau + \delta)$ to allow for their loss of the option to hold physical currency over the period $t + \tau$ to $t + \tau + \delta$. From a borrower's perspective, they will be prepared to offer the compensation $C(t, \tau, \tau + \delta)$ to the investor because the receipts $\underline{P}(t, \tau, \tau + \delta)$ from the forward bond sale at time $t + \tau$ also come with the option to invest in physical currency if that offers a higher return than the shadow term structure at time $t + \tau$. Options to hold physical currency cannot, of course, be traded independently of the forward bonds; the options are only available to those with surplus funds in the current finite-step or projected surplus funds in future finite-steps.

In explicit notation, the valuation of a finite-step ZLB forward bond is analogous to the shadow finite-step set-up in table 1, but I replace the shadow forward bond price $P(t, \tau, \tau + \delta)$ with the ZLB forward bond price $\underline{P}(t, \tau, \tau + \delta)$ and add an additional column to account for the option price $C(t, \tau, \tau + \delta)$. Note that separating the discount factor $\exp\left(-\int_t^{t+\tau} r(v) dv\right)$ from the contingent cashflow $\max[P(t + \tau, \delta) - 1, 0]$ is arguably an abuse of notation, but it serves to retain the analogy with section 2.1.1. In section 3, I provide a more formal framework and also explain why the shadow term structure (rather than the ZLB term structure) provides the appropriate discount factors for each of the cashflows.

Table 2: Forward bond with physical currency			
Time	$t + \tau$	$t + \tau + \delta$	$t + \tau$
Cashflows	$-\underline{P}(t, \tau, \tau + \delta)$	1	$\max[P(t + \tau, \delta) - 1, 0]$
Discount factors	$P(t, \tau)$	$P(t, \tau + \delta)$	$\exp\left(-\int_0^\tau r(v) dv\right)$
Present values	$-P(t, \tau) \cdot \underline{P}(t, \tau, \tau + \delta)$	$P(t, \tau + \delta)$	$C(t, \tau, \tau + \delta)$

The set-up in table 2 results in the arbitrage-free expression for the ZLB forward bond in terms of bond and option prices for the shadow term structure; i.e. $-P(t, \tau) \cdot \underline{P}(t, \tau, \tau + \delta) + P(t, \tau + \delta) = C(t, \tau, \tau + \delta)$, and so:

$$\underline{P}(t, \tau, \tau + \delta) = \frac{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)}{P(t, \tau)} \quad (14)$$

The associated finite-step ZLB forward rate $\underline{F}(t, \tau, \tau + \delta)$ is:

$$\begin{aligned} \underline{F}(t, \tau, \tau + \delta) &= -\frac{1}{\delta} \log [\underline{P}(t, \tau, \tau + \delta)] \\ &= -\frac{1}{\delta} \log \left[\frac{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \end{aligned} \quad (15)$$

2.3.2 Continuous-time ZLB forward rate

Analogous to calculating the continuous-time shadow forward rate $f(t, \tau)$, the continuous-time ZLB forward rate $\underline{f}(t, \tau)$ may be evaluated as the infinitesimal limit of the finite-step ZLB forward rate $\underline{F}(t, \tau, \tau + \delta)$. I have relegated the details to appendix A because the derivation is longer than for the shadow forward rate $f(t, \tau)$ in section 2.1.2. However, it is worthwhile highlighting here that the calculations remain straightforward, requiring just analytic (rather than stochastic) calculus because only analytic expressions are involved. To summarize the results:

$$\begin{aligned} \underline{f}(t, \tau) &= \lim_{\delta \rightarrow 0} \underline{F}(t, \tau, \tau + \delta) \\ &= f(t, \tau) + \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right\} \\ &= f(t, \tau) + z(t, \tau) \end{aligned} \quad (16)$$

where $z(t, \tau)$ is the option-adjustment term:

$$z(t, \tau) = \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right\} \quad (17)$$

which will depend on the specific model used to represent shadow term structure.

The instantaneous ZLB forward rate $\underline{f}(t, \tau)$ is therefore the shadow forward rate $f(t, \tau)$ plus an option-adjustment term $z(t, \tau)$ that explicitly accounts for the effect of the availability of physical currency on the term structure. The ZLB forward rate $\underline{f}(t, \tau)$ establishes the ZLB interest rate and bond price expressions in the usual way. i.e.:

$$\underline{r}(t, \tau) = \frac{1}{\tau} \int_0^\tau \underline{f}(t, v) dv \quad (18)$$

and:

$$\underline{P}(t, \tau) = \exp[-\tau \cdot \underline{r}(t, \tau)] \quad (19)$$

3 ZLB framework for the generic GATSM

In this section, I develop the ZLB-GATSM framework using the concepts outlined in section 2.3 but explicitly using the generic GATSM (i.e. with an arbitrary number N factors, factor inter-relationships, and innovation correlations) to represent the shadow term structure. Section 3.1 outlines the specification and results for the generic GATSM relevant to the present article. Section 3.2 contains my original contribution, first using the generic GATSM specification to provide a formal specification and derivation for the results outlined in section 2.3, and then deriving the generic ZLB-GATSM based on the generic GATSM expressions.

Note that I assume all standard technical conditions regarding probability spaces etc., and (mainly for notional clarity) all quantities and expectations I use to develop my ZLB framework are under the risk-neutral or \mathbb{Q} measure (i.e. with zero market prices of risk for each factor, so without risk premiums in the associated term structure). I discuss the extension to the physical or \mathbb{P} measure (i.e. with non-zero market prices of risk, thereby allowing for risk premiums in the associated term structure) in section 5.¹¹

3.1 The generic GATSM term structure

Section 3.1.1 first outlines the general equilibrium economy foundation for the generic GATSM, the resulting stochastic process for the short rate, and the associated partial differential equation. Sections 3.1.2 to 3.1.3 respectively provide the generic GATSM expressions for bond prices and option prices. While GATSM summaries are readily available in many articles and textbooks, I generally use Chen (1995) as a convenient point of reference because it contains explicit multifactor expressions for GATSM bond and option prices.¹² However, I also refer to Berardi and Esposito (1999) and Berardi (2009) for their economic foundation,¹³ and Dai and Singleton (2002) pp. 437-8 for

¹¹Filipović (2009) chapter 4 contains further details on these aspects.

¹²Chen (1995) actually provides the specification and results for a two-factor GATSM for bond and option prices, but I allow for the ready extension to N factors that Chen (1995) notes. Chaplin (1987) first derived explicit bond and option prices for the two-factor GATSM, and Sharp (1987) extended those expressions to N factors. The heritage for multifactor GATSMs appears to begin with Langetieg (1980).

¹³Duffie (1996) is a useful textbook reference with extensive coverage of the equilibrium foundations for dynamic asset pricing.

the convenience of its matrix notation. The GATSM specifications in these articles are equivalent, given they are all within affine-invariant transformations of each other.¹⁴

3.1.1 A general equilibrium economy foundation for GATSMs

Berardi and Esposito (1999, hereafter BE) provides a standard general equilibrium basis for multifactor GATSMs based on a Cox, Ingersoll, and Ross (1985a, hereafter CIR) representative-agent economy. The overview of the BE/CIR economy is best summarized as a direct quote from BE p. 147:

“The theoretical framework is based on the maximizing behavior of a fixed number of identical individuals with rational expectations and time-additive logarithmic utility functions. The economy is competitive, with continuous trading and no transactions costs, and it is characterized by the existence of markets for contingent claims and for instantaneous borrowing and lending at the riskless interest rate. In the economy, a single physical good is produced, which may be allocated to consumption or investment.¹⁵ The investment opportunities consist of a stochastic production process, a set of contingent claims and a risk-free asset. In equilibrium, the rate of return on these securities must be such that the wealth of individuals is totally invested in the production process. The equilibrium of the economy is influenced by a vector of $[N - 1]$ unknown state variables” “As in [CIR], we assume that the mean of the rates of return on the production process are proportional to the state variables.”

As discussed in Berardi (2009) pp. 990-2, the $N - 1$ state variables may be interpreted as representing the returns to the factors of production (e.g. land, capital, labor, etc.) that determine the real term structure. BE pp. 151-55 also follows the arguments of Cox, Ingersoll, and Ross (1985b) to introduce a single state variable representing inflation on consumption goods along with the CIR assumption of the neutrality of money (i.e. changes in the money supply have no real effects). BE does not specify an explicit monetary policy process for the BE/CIR economy, but Berardi (2009) pp. 990-2 assumes an underlying equilibrium in the money market that supports an exogenously given process for the price level, and also discusses how the long-run levels (μ_n in equation 21 below) of the real state variables and the inflation state variable effectively proxy the targets for the central bank.¹⁶

Using the notation from Chen (1995), the BE/CIR economy in its purely Gaussian form leads to the following expression for the generic GATSM nominal short rate:¹⁷

$$r(t) = \sum_{n=1}^N s_n(t) \quad (20)$$

¹⁴See Singleton (2006) pp. 319-21 for details.

¹⁵As in CIR, all values are expressed in terms of units of this good.

¹⁶Indeed, Berardi (2009) allows for Gaussian innovations to μ_n so the targets can vary over time, but the model remains a GATSM.

¹⁷BE actually specifies N real state variables, but respecifying as a total of $N - 1$ real state variables plus a single inflation state variable for a total of N state variables is notationally more convenient.

where $r(t)$ is the instantaneous short rate at time t , and $s_n(t)$ are the N state variables which evolve as a correlated Ornstein-Uhlenbeck process:

$$ds_n(t) = \kappa_n [\mu_n - s_n(t)] + \sigma_n dW_n(t) \quad (21)$$

where μ_n are constants representing the long-run levels of $s_n(t)$, κ_n are positive constants representing the mean reversion rates of $s_n(t)$ to μ_n , σ_n are positive constants representing the volatilities (annualized standard deviations) of $s_n(t)$, and $W_n(t)$ are Wiener processes with $dW_n(t) \sim N(0, 1)dt$ and $\mathbb{E}[dW_m(t), dW_n(t)] = \rho_{mn}dt$, where ρ_{mn} are correlations $-1 \leq \rho_{mn} \leq 1$.

Again using Chen (1995) notation, the BE/CIR model economy results in the following partial differential equation (hereafter PDE) for the valuation of interest rate instruments:¹⁸

$$\left[\sum_{n=1}^N s_n(t) \right] \cdot X(\cdot) = \frac{\partial X(\cdot)}{\partial t} + \sum_{n=1}^N \frac{\partial X(\cdot)}{\partial s_n(t)} \kappa_n [\mu_n - s_n(t)] + \frac{1}{2} \sum_{n=1}^N \frac{\partial^2 X(\cdot)}{\partial [s_n(t)]^2} \sigma_n^2 + \sum_{m=1}^N \sum_{n=1}^N \frac{\partial^2 X(\cdot)}{\partial s_m(t) \partial s_n(t)} \rho_{mn} \sigma_m \sigma_n \quad (22)$$

The PDE may be used to obtain solutions for interest rate instruments $X(\cdot)$ dependent on the nature of their boundary conditions:

- The boundary condition $P(t + \tau, 0) = 1$ produces the solution for the bond $P(t, \tau)$, as subsequently given in section 3.1.2.
- The boundary condition $C(t + \tau, 0, \delta) = \max\{0, P(t + \tau, \delta) - 1\}$ produces the solution for the call option $C(t, \tau, \tau + \delta)$, as subsequently given in section 3.1.3.
- The boundary condition $Q(t + \tau, 0, \delta) = \max\{0, 1 - P(t + \tau, \delta)\}$ produces the solution for the put option $Q(t, \tau, \tau + \delta)$.¹⁹ An explicit expression for $Q(t, \tau, \tau + \delta)$ is not actually required, but introducing it simplifies my subsequent ZLB forward rate derivation in section 3.2.1.

3.1.2 GATSM bond prices

Chen (1995) p. 349-50 provides the following closed-form analytic expression for GATSM bond prices:²⁰

$$P(t, \tau) = \exp \left[-H(\tau) - \sum_{n=1}^N s_n(t) \cdot G(\kappa_n, \tau) \right] \quad (23)$$

¹⁸The BE/CIR specification renormalizes the original state variables $x_n(t)$ to obtain orthogonal state variables $s_n(t)$ so that $\rho_{mn} = 0$ and the double-summation expression evaluates to zero. I have retained the more flexible specification from Chen (1995) pp. 349-50 that allows for non-zero innovation correlations.

¹⁹Expressions for put option prices are often denoted $P(\cdot)$, but I have used the notation $Q(\cdot)$ to avoid any confusion with my notation $P(\cdot)$ for bond prices. Of course, $Q(\cdot)$ has no connection to the risk-neutral \mathbb{Q} measure.

²⁰Chen (1995) actually uses time t and time of maturity T notation, with the PDE boundary condition $P(T, T) = 1$ to obtain an expression for $P(t, T)$. My time t and time to maturity $\tau = T - t$ notation is equivalent.

where the functions $G(\kappa_n, \tau)$ are:

$$G(\kappa_n, \tau) = \frac{1}{\kappa_n} [1 - \exp(-\kappa_n \tau)] \quad (24)$$

and the function $H(\tau)$ for the generic GATSM is most concisely expressed using Dai and Singleton (2002) matrix notation:²¹

$$H(\tau) = -\frac{1}{2} \text{Tr} [\tau \Xi(\tau) \Psi] + \sum_{n=1}^N \mu_n [\tau - G(\kappa_n, \tau)] \quad (25)$$

where $\text{Tr}[\cdot]$ is the matrix trace operator, the matrix Ψ is:

$$\Psi_{ij} = \frac{1}{\kappa_i \kappa_j} \quad (26)$$

and the matrix $\tau \Xi(\tau)$ is:

$$\tau \Xi_{ij}(\tau) = \rho_{ij} \sigma_i \sigma_j \cdot [\tau - G(\kappa_i, \tau) - G(\kappa_j, \tau) + G(\kappa_i + \kappa_j, \tau)] \quad (27)$$

3.1.3 GATSM call option prices

Chen (1995) p. 360 provides the following closed-form analytic expressions for GATSM option prices:²²

$$C(t, \tau, \tau + \delta) = P(t, \tau + \delta) N[d_1(t, \tau, \tau + \delta)] - P(t, \tau) N[d_2(t, \tau, \tau + \delta)] \quad (28)$$

where:

$$d_1(t, \tau, \tau + \delta) = \frac{1}{\Sigma(\tau, \tau + \delta)} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] + \frac{1}{2} \Sigma(\tau, \tau + \delta) \quad (29)$$

$$d_2(t, \tau, \tau + \delta) = d_1(t, \tau, \tau + \delta) - \Sigma(\tau, \tau + \delta) \quad (30)$$

and:²³

$$\begin{aligned} [\Sigma(\tau, \tau + \delta)]^2 &= \sum_{m=1}^N \sigma_m^2 [G(\kappa_m, \delta)]^2 G(2\kappa_m, \tau) \\ &+ 2 \sum_{m=1}^N \sum_{n=m+1}^N \rho_{mn} \sigma_m \sigma_n G(\kappa_m, \delta) G(\kappa_n, \delta) G(\kappa_m + \kappa_n, \tau) \end{aligned} \quad (31)$$

²¹The expressions for $H(\tau)$ and $\tau \Xi_{ij}(\tau)$ arise from substituting the quantities $X = I$, $b_0 = [1, \dots, 1]'$, $a_0 = 0$, $\theta = 0$, $\kappa = \text{diag}[\kappa_1, \dots, \kappa_N]$, and $\Sigma = \rho_{ij} \sigma_i \sigma_j$ into the Dai and Singleton (2002) specification. Chen (1995) gives the two-factor result in summation form and Vincente and Tabak (2008) gives the N -factor result in double-summation form.

²²Chen (1995) actually uses time t , time of option expiry T_C , and time of maturity T notation, with a strike price K . Setting $T_C = t + \tau$, $T = t + \tau + \delta$, and $K = 1$ produces my equivalent expressions.

²³See Chen (1995) p. 351 for the volatility expression $\Sigma(\tau, \tau + \delta)$, and p. 348 for the variance and covariance expressions used in $\Sigma(\tau, \tau + \delta)$. I express the variances and covariances more conveniently in terms of $G(\cdot, \cdot)$ functions.

3.2 The generic ZLB-GATSM term structure

This section contains my original contribution developing the ZLB-GATSM framework. Using the BE/CIR economy from section 3.1.1, I first provide a formal general equilibrium economy foundation for the generic ZLB-GATSM, justifying the ZLB forward rate expression introduced in section 2.3. Note that while the justification is specifically for the ZLB-GATSM, I discuss in section 5 how the results may readily be extended to accommodate term structure models with Cox, Ingersoll, and Ross (1985b)/square-root dynamics.

Section 3.2.2 derives the shadow forward rate and section 3.2.3 derives the option-adjustment term. While I relegate the details to appendix B, it is worthwhile noting again that the calculations only require analytic (rather than stochastic) calculus because only analytic expressions are involved. The resulting forward rate and option-adjustment term are combined to obtain the generic ZLB-GATSM forward rate expression in section 3.2.3, and then I make a series of observations on ZLB-GATSM forward rate in section 3.2.4.

3.2.1 A general equilibrium foundation for ZLB-GATSMs

I start by assuming that the BE/CIR economy and its general equilibrium remains invariant to physical currency being made available in the economy. That invariance assumption is crucial to the derivation of closed-form ZLB-GATSM forward rates and it also might appear to be overly strong on first reading, so I spend the following two paragraphs justifying my assumption.

Regarding the real component of the BE/CIR economy, the BE/CIR neutrality of money assumption is sufficient to support my assumption that the real component of the BE/CIR economy remains invariant to the availability of physical currency. However, my assumed invariance would also stand even allowing for interactions between the real economy and money (as discussed by BE p. 151) so long as the physical currency component of money is not regarded as a real factor of production. That follows from the BE/CIR assumption that “*In equilibrium, ... the wealth of individuals is totally invested in the production process*”. Finally, even if plausible models could be specified giving physical currency an explicit role in the production process, the empirical evidence suggests the role is small. For example, U.S. Federal Reserve figures for 2011:Q2 indicate that physical currency comprised less than 1.7 percent of the total net physical and financial assets (i.e. excluding human capital and other intangible assets) of households in the U.S.²⁴

Regarding the inflation component of the BE/CIR economy, the BE/CIR assumption of an exogenous process for the price level is sufficient to ensure that inflation remains invariant to the availability of physical currency. However, my assumed invariance would still stand even allowing for some endogeneity between the real economy and the price level because the availability of physical currency is not, in its own right, a singular influence on inflation. To clarify my point, I make the distinction between the availability of physical currency per se, and monetary policy. It is monetary policy and its operation (e.g. specifying and targeting nominal quantities, such as growth rates

²⁴That is, a \$976.6 billion average currency component from monthly H.6 Table 2 M1 versus \$58.73 trillion from the B.100 Balance Sheet of Households and Nonprofit Organizations.

for the monetary base or inflation rates directly) that are the underlying determinants of inflation.

With my assumed invariance of the BE/CIR economy (or similar models) to the availability of physical currency, the shadow short rate remains the sum of the BE/CIR state variables (as in equation 20) and the state variables themselves retain their GATSM dynamics (as in equation 21). Most importantly, the PDE for the valuation of interest rate instruments in the BE/CIR economy remains unchanged from equation 22.

What does change in the BE/CIR economy is the prices of the securities that fund the investment in the production process because, repeating the already-noted assumption used for the BE/CIR economy, “*In equilibrium, the rate of return on these securities must be such that the wealth of individuals is totally invested in the production process*”. Therefore, to avoid any incentive for investors to hold physical currency rather than investing in the production process, bond prices must change so that they offer a rate of return that compensates investors for the optionality associated with the availability of physical currency. The required bond price may be obtained by using the boundary condition for the ZLB bond in the PDE for the BE/CIR economy; i.e. $\underline{P}(t + \tau, \delta) = \min \{1, P(t + \tau, \delta)\}$. That boundary condition may be re-expressed as the boundary condition for a shadow bond, i.e. 1, less the boundary condition for a put option, i.e.:

$$\begin{aligned} \underline{P}(t + \tau, \delta) &= \min \{1, P(t + \tau, \delta)\} \\ &= 1 + \min \{0, P(t + \tau, \delta) - 1\} \\ &= 1 - \max \{0, 1 - P(t + \tau, \delta)\} \end{aligned} \tag{32}$$

Therefore, the solution for $\underline{P}(t, \tau + \delta)$ must be the shadow bond price less the put option price for the shadow term structure, i.e.:

$$\underline{P}(t, \tau + \delta) = P(t, \tau) - Q(t, \tau, \tau + \delta) \tag{33}$$

Using the standard put-call parity relationship,²⁵ equation 33 may be re-expressed in terms of the $\tau + \delta$ -maturity bond and the call option, i.e:

$$\begin{aligned} C(t, \tau, \tau + \delta) + P(t, \tau) &= Q(t, \tau, \tau + \delta) + P(t, \tau + \delta) \\ P(t, \tau) - Q(t, \tau, \tau + \delta) &= P(t, \tau + \delta) - C(t, \tau, \tau + \delta) \\ \underline{P}(t, \tau + \delta) &= P(t, \tau + \delta) - C(t, \tau, \tau + \delta) \end{aligned} \tag{34}$$

The ZLB forward rate $\underline{f}(t, \tau)$ associated with the ZLB bond price $\underline{P}(t, \tau + \delta)$ may be calculated directly using the definition of the forward rate with respect to the bond

²⁵That is, the forward price equals the call option price less the put option price, or $F = C - P$. See Haug (2007) p. 18 or the original reference therein to Nelson (1904).

price, and then taking the limit, i.e:

$$\begin{aligned}
\underline{f}(t, \tau) &= \lim_{\delta \rightarrow 0} \underline{f}(t, \tau + \delta) \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d[\tau + \delta]} \log [\underline{P}(t, \tau + \delta)] \right\} \\
\langle \text{Chain rule} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\tau} \log [\underline{P}(t, \tau + \delta)] \cdot \frac{d\tau}{d[\tau + \delta]} \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\tau} \log [\underline{P}(t, \tau + \delta)] \right\} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d\tau}{d[\tau + \delta]} \right\} \\
\langle \text{Limit rules} \rangle &= \frac{d}{d\tau} \log [\underline{P}(t, \tau)] \cdot \frac{1}{\lim_{\delta \rightarrow 0} \left\{ \frac{d[\tau + \delta]}{d\tau} \right\}} \\
&= [f(t, \tau) + z(t, \tau)] \cdot \frac{1}{\lim_{\delta \rightarrow 0} \{1\}} \\
&= f(t, \tau) + z(t, \tau)
\end{aligned} \tag{35}$$

where ‘‘Limit rules’’ refers to the standard rules for manipulating limits.

Equation 35 is identical to equation 16, but with a formal foundation in the BE/CIR general equilibrium economy. It is now also apparent why the shadow short rate and the associated shadow term structure were used to provide the appropriate discount factors for the present-value set-up in table 2 (i.e. the shadow bond prices $\underline{P}(t, \tau)$ and $\underline{P}(t, \tau + \delta)$ for the cashflows $\underline{P}(t, \tau, \tau + \delta)$ and 1, and the expected path of the shadow short rate $r(t)$ for the option price $C(t, \tau, \tau + \delta)$).

The explicit expression for the generic ZLB-GATSM forward rate $\underline{f}(t, \tau)$ simply requires explicit expressions for the shadow/GATSM forward rate $f(t, \tau)$ and the option-adjustment expression $z(t, \tau)$. These are derived in the following two subsections.

3.2.2 Shadow/GATSM forward rates

GATSM forward rates may be obtained using the standard bond price/forward rate definition and the closed-form analytic expression for generic GATSM bond prices from section 3.1.2. Appendix B.1 contains the details, and the resulting expression for the generic GATSM forward rate is:

$$\begin{aligned}
f(t, \tau) &= -\frac{d}{d\tau} \log [\underline{P}(t, \tau)] \\
&= -\frac{1}{2} \text{Tr} [\Theta(\tau) \Psi] + \sum_{n=1}^N \mu_n + \sum_{n=1}^N [s_n(t) - \mu_n] \exp(-\kappa_n \tau)
\end{aligned} \tag{36}$$

where the matrix $\Theta(\tau)$ is:

$$\Theta_{ij}(\tau) = \rho_{ij} \sigma_i \sigma_j \cdot \kappa_i \kappa_j G(\kappa_i, \tau) G(\kappa_j, \tau) \tag{37}$$

3.2.3 The option-adjustment expression

The option-adjustment term $z(t, \tau)$ may be obtained using the closed-form analytic expressions for GATSM bond and call option prices. Appendix B.2 contains the details,

and the resulting expression for the generic GATSM option-adjustment term is:

$$\begin{aligned} z(t, \tau) &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right\} \\ &= -f(t, \tau) \cdot \left(1 - N \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \end{aligned} \quad (38)$$

where $\omega(\tau)$ is the annualized instantaneous volatility:

$$\begin{aligned} \omega(\tau) &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\delta} \Sigma(\tau, \tau + \delta) \right\} \\ &= \sqrt{\sum_{n=1}^N \sigma_n^2 \cdot [G(2\kappa_n, \tau)]^2 + 2 \sum_{m=1}^N \sum_{n=m+1}^N \rho_{mn} \sigma_m \sigma_n \cdot G(\kappa_m + \kappa_n, \tau)} \end{aligned} \quad (39)$$

3.2.4 ZLB-GATSM forward rates

Substituting the result for $z(t, \tau)$ from equation 38 into equation 16 gives the generic ZLB-GATSM forward rate expression:

$$\begin{aligned} \underline{f}(t, \tau) &= f(t, \tau) + z(t, \tau) \\ &= f(t, \tau) - f(t, \tau) \left(1 - N \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \\ &= f(t, \tau) \cdot N \left[\frac{f(t, \tau)}{\omega(\tau)} \right] + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \end{aligned} \quad (40)$$

The ZLB-GATSM interest rate expression $\underline{r}(t, \tau)$ and bond price expression $\underline{P}(t, \tau)$ have already been given respectively in equations 18 and 19. Specifically, $\underline{r}(t, \tau)$ may be calculated for any given time to maturity by numerically integrating $\underline{f}(t, \tau)$ between the limits 0 to τ ,²⁶ and $\underline{P}(t, \tau)$ is then a scalar exponential involving $\underline{r}(t, \tau)$.

3.2.5 Observations on the generic ZLB-GATSM framework

One advantage of developing the ZLB-GATSM framework on a completely generic basis is that the associated observations will apply to any particular ZLB-GATSM specification, regardless of the number of factors, factor inter-relationships, and innovation correlations. As a related point, the first observation is that the ZLB-GATSM framework obviously preserves the complete flexibility of the GATSM class of models, because any GATSM may be used to represent the shadow term structure and the ZLB-GATSM framework then provides the required modification to ensure the ZLB is respected in a theoretically consistent manner.

The second observation is that ZLB-GATSM forward rates $\underline{f}(t, \tau)$ will always have simple closed-form analytic expressions. That is evident from the generic ZLB-GATSM

²⁶I have tried, obviously without success, to derive closed-form analytic expressions for $\underline{r}(t, \tau)$, and hence $\underline{P}(t, \tau)$. The form of $\underline{f}(t, \tau)$ tantalizingly suggests the possibility of an analytic integral, but integration by parts did not prove successful and neither did “brute force” analytic integration via Mathematica.

being itself composed of simple closed-form analytic expressions, i.e.: (1) $f(t, \tau)$, which is composed of scalar exponential functions of time to maturity τ and the state variables; (2) $\omega(\tau)$, which is composed of scalar exponential functions of τ and the state variable innovation variances and covariances; (3) the univariate cumulative normal distribution $N[f(t, \tau) / \omega(\tau)]$, which is a standard function that is well tabulated or readily approximated with closed-form analytic expressions; and (4) the scalar exponential function $\exp(-\frac{1}{2} [f(t, \tau) / \omega(\tau)]^2)$.

Third, the numerical integration required to obtain $\underline{r}(t, \tau)$ does not change in complexity as the specification of the ZLB-GATSM is changed, because the functions required for the numerical integration, i.e. $N[f(t, \tau) / \omega(\tau)]$ and $\exp(-\frac{1}{2} [f(t, \tau) / \omega(\tau)]^2)$, always remain univariate. The inputs to those functions, i.e. $f(t, \tau)$ and $\omega(\tau)$, are always closed-form analytic expressions that are calculated prior to the cumulative normal and exponential evaluations. The guarantee of invariant numerical complexity regardless of specification means that ZLB-GATSMs retain a large degree of tractability for empirical applications and estimations. By comparison, univariate numerical integration is certainly more straightforward than the multivariate numerical methods used to apply the Black (1995) approach to the two-factor GATSMs mentioned in the introduction (i.e. Monte Carlo, interest rate lattice, and finite-difference methods).²⁷

Fourth, the ZLB-GATSM forward rates converge to GATSM forward rates when the latter are sufficiently positive relative to term structure volatility. Specifically, as the ratio $f(t, \tau) / \omega(\tau)$ increases to larger positive values, $N[f(t, \tau) / \omega(\tau)]$ approaches one and $\exp(-\frac{1}{2} [f(t, \tau) / \omega(\tau)]^2)$ approaches zero, so $\lim_{f(t, \tau) / \omega(\tau) \rightarrow \infty} \underline{f}(t, \tau) = f(t, \tau)$. The convergence of ZLB and shadow/GATSM forward rates functions means that ZLB interest rates and bond prices also converge to their shadow counterparts.

Fifth and finally, the ZLB-GATSM must respect the lower bound of zero no matter how negative shadow forward rates become. Specifically, as the ratio $f(t, \tau) / \omega(\tau)$ decreases to larger negative values, $N[f(t, \tau) / \omega(\tau)]$ and $\exp(-\frac{1}{2} [f(t, \tau) / \omega(\tau)]^2)$ both approach zero, so $\lim_{f(t, \tau) / \omega(\tau) \rightarrow -\infty} \underline{f}(t, \tau) = 0$.

4 Examples to illustrate the ZLB-GATSM framework

In this section, I illustrate the salient points the ZLB-GATSM framework using a calibrated risk-neutral Heath, Jarrow, and Morton (1992, pp. 90-92, hereafter HJM) model to represent the shadow term structure. The HJM model, which I derive in section 4.1 as a special case of the unrestricted two-factor GATSM, is ideal for the purposes of demonstration for several reasons: (1) it is parsimonious enough to facilitate simple and transparent examples, while two factors is flexible enough to produce plausible term structures and emphasize that the ZLB-GATSM framework readily accommodates multifactor GATSM models; (2) it contains unit root dynamics that result in negative forward rates at very long horizons, thus providing an illustration of how the ZLB-GATSM framework resolves that issue; and (3) it is closely related to the Christensen, Diebold, and Rudebusch (2011, hereafter CDR) arbitrage-free version of

²⁷Indeed, even the Gorovoi and Linetsky (2004) analytical expression for ZLB bond prices in the single-factor GATSM or Vasicek (1977) model is not closed form, and the numerical evaluation of relatively complex functions (e.g. Weber-Hermite parabolic cylinder functions) is required.

the Nelson and Siegel (1987) model, so I can adopt their parameter estimates for my illustrations. Regarding the latter point, I reiterate that the aim of this section is purely an illustration in the context of this article having introduced the ZLB-GATSM framework; section 5 discusses why space limitations would make a complete empirical estimation of the ZLB-GATSM framework beyond the scope of this article (but I intend to undertake such an estimation in future research).

Section 4.2 focusses on the cross-sectional perspective of the ZLB-HJM model. The first two examples illustrate how, at a given point in time, the ZLB-HJM model transforms negative shadow forward rates and interest rates for short horizons/times to maturity into non-negative values. The third example illustrates the transformation for very long horizons/times to maturity. Section 4.3 focusses on the time series perspective of the ZLB-HJM model, showing the evolution of the state variables estimated from observations of term structure data.

4.1 The HJM model within the ZLB-GATSM framework

The HJM model is a two-factor GATSM with long-run levels of zero for the state variables, uncorrelated factor innovations, and a zero rate of mean reversion for one of the factors. Appendix C details how the HJM model may be replicated within the generic GATSM framework. The resulting forward rate expression, with the non-zero rate of mean reversion set to λ for notational convenience, is:

$$f(t, \tau) = s_1(t) + s_2(t) \cdot \exp(-\lambda\tau) - \sigma_1^2 \cdot \frac{1}{2}\tau^2 - \sigma_2^2 \cdot \frac{1}{2} [G(\lambda, \tau)]^2 \quad (41)$$

and the annualized instantaneous volatility is:

$$\omega(\tau) = \sqrt{\sigma_1^2 \cdot \tau^2 + \sigma_2^2 \cdot [G(2\lambda, \tau)]^2} \quad (42)$$

The forward rate factor loading for the state variable $s_1(t)$ is therefore constant at 1 for all times to maturity, so $s_1(t) \cdot 1$ can be viewed as the “Level” component of the HJM term structure. The forward rate factor loading for the state variable $s_2(t)$ is an exponential decay as a function of time to maturity, so $s_2(t) \cdot \exp(-\lambda\tau)$ can be viewed as the “Slope” component of the HJM term structure.

The HJM model may also be viewed as the CDR “Level, Slope, Curvature” model without the Curvature component. I therefore use the parameters related to Level and Slope from CDR (p. 12 table 5, i.e. $\lambda = 0.5975$, $\sigma_1 = 0.0051$, and $\sigma_2 = 0.0110$) to calibrate my HJM model.

The ZLB-HJM forward rate expression is then:

$$\underline{f}(t, \tau) = f(t, \tau) \cdot N \left[\frac{f(t, \tau)}{\omega(\tau)} \right] + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \quad (43)$$

and the ZLB-HJM interest rate for any given time to maturity τ is $\underline{r}(t, \tau) = \frac{1}{\tau} \int_0^\tau \underline{f}(t, v) dv$. I calculate the latter by numerical integration using the rectangular/mid-ordinate rule, i.e.:

$$\underline{r}(t, \tau) = \frac{1}{\tau} \left[\Delta\tau \sum_{i=1}^I \underline{f} \left(t, \left[i - \frac{1}{2} \right] \Delta\tau \right) \right] \quad (44)$$

Equation 44 conveniently turns into a simple mean, because $I = \frac{\tau}{\Delta\tau}$, so $\Delta\tau = \frac{\tau}{I}$, and then $\frac{1}{\tau}\Delta\tau = \frac{1}{\tau I} = \frac{1}{I}$. Therefore:

$$\begin{aligned} \mathbf{r}(t, \tau) &= \frac{1}{I} \sum_{i=1}^I \mathbf{f}\left(t, \left[i - \frac{1}{2}\right] \Delta\tau\right) \\ &= \text{mean}\{\mathbf{f}(t, 0.5\Delta\tau), \dots, \mathbf{f}(t, \tau - 0.5\Delta\tau)\} \end{aligned} \quad (45)$$

and I ensure all integrals are accurate to within 5e-8, which is 5/10000th of a basis point.²⁸

The shadow interest rate expression may be obtained by applying the transformation $\mathbf{r}(t, \tau) = \frac{1}{\tau} \int_0^\tau \mathbf{f}(t, v) dv$ to each component of the shadow forward rate expression, and the results are closed-form analytic expressions. That is: $\frac{1}{\tau} \int_0^\tau s_1(t) dv = s_1(t)$, $\frac{1}{\tau} \int_0^\tau s_2(t) \cdot \exp(-\lambda\tau) dv = s_2(t) \cdot \frac{1 - \exp(-\lambda\tau)}{\lambda\tau} = s_2(t) \cdot \frac{1}{\tau} G(\lambda, \tau)$, $\frac{1}{\tau} \int_0^\tau \sigma_1^2 \cdot \frac{1}{2} v^2 dv = \sigma_1^2 \cdot \frac{1}{6} \tau^2$, $\frac{1}{\tau} \int_0^\tau \sigma_2^2 \cdot \frac{1}{2} [G(\lambda, v)]^2 dv = \sigma_2^2 \cdot \left[\frac{1}{2\lambda^2} + \frac{1}{\lambda^3} \frac{1 - \exp(-\lambda\tau)}{\tau} + \frac{1}{4\lambda^3} \frac{1 - \exp(-2\lambda\tau)}{\tau} \right] = \sigma_2^2 \cdot \frac{1}{\lambda^2} \left[\frac{1}{2} + \frac{1}{\tau} G(\lambda, \tau) + \frac{1}{2\tau} G(2\lambda, \tau) \right]$, and so the shadow interest rate expression is therefore:²⁹

$$\begin{aligned} \mathbf{r}(t, \tau) &= s_1(t) + s_2(t) \cdot \frac{1}{\tau} G(\lambda, \tau) \\ &\quad - \sigma_1^2 \cdot \frac{1}{6} \tau^2 - \sigma_2^2 \cdot \frac{1}{\lambda^2} \left[\frac{1}{2} + \frac{1}{\tau} G(\lambda, \tau) + \frac{1}{2\tau} G(2\lambda, \tau) \right] \end{aligned} \quad (46)$$

4.2 Cross-sectional/time-to-maturity perspective

Figure 1 shows how the ZLB-GATSM framework accommodates a materially negative shadow short rate and its associated term structure. Specifically, I set the ZLB-HJM model state variables to $s_1(t) = 0.05$ and $s_2(t) = -0.10$, which gives a shadow short rate value of $\mathbf{r}(t) = s_1(t) + s_2(t) = -0.05$, or -5 percent.

²⁸One percentage point is 0.01 and one basis point is 0.01 of a percentage point, so 5e-8 = 5e-4 basis points. Note that 5e-8 is an arbitrary choice, but is certainly small enough to ensure the examples will not be unduly influenced by the method of numerical integration. A value of $\Delta\tau = 0.00125$ years proved suitable for obtaining my 5e-8 accuracy threshold in all of the examples presented in this article.

²⁹These results accord with those available in CDR pp. 6-7, but using just their Level and Slope factors and setting their $\sigma_{12} = \sigma_{21} = 0$.

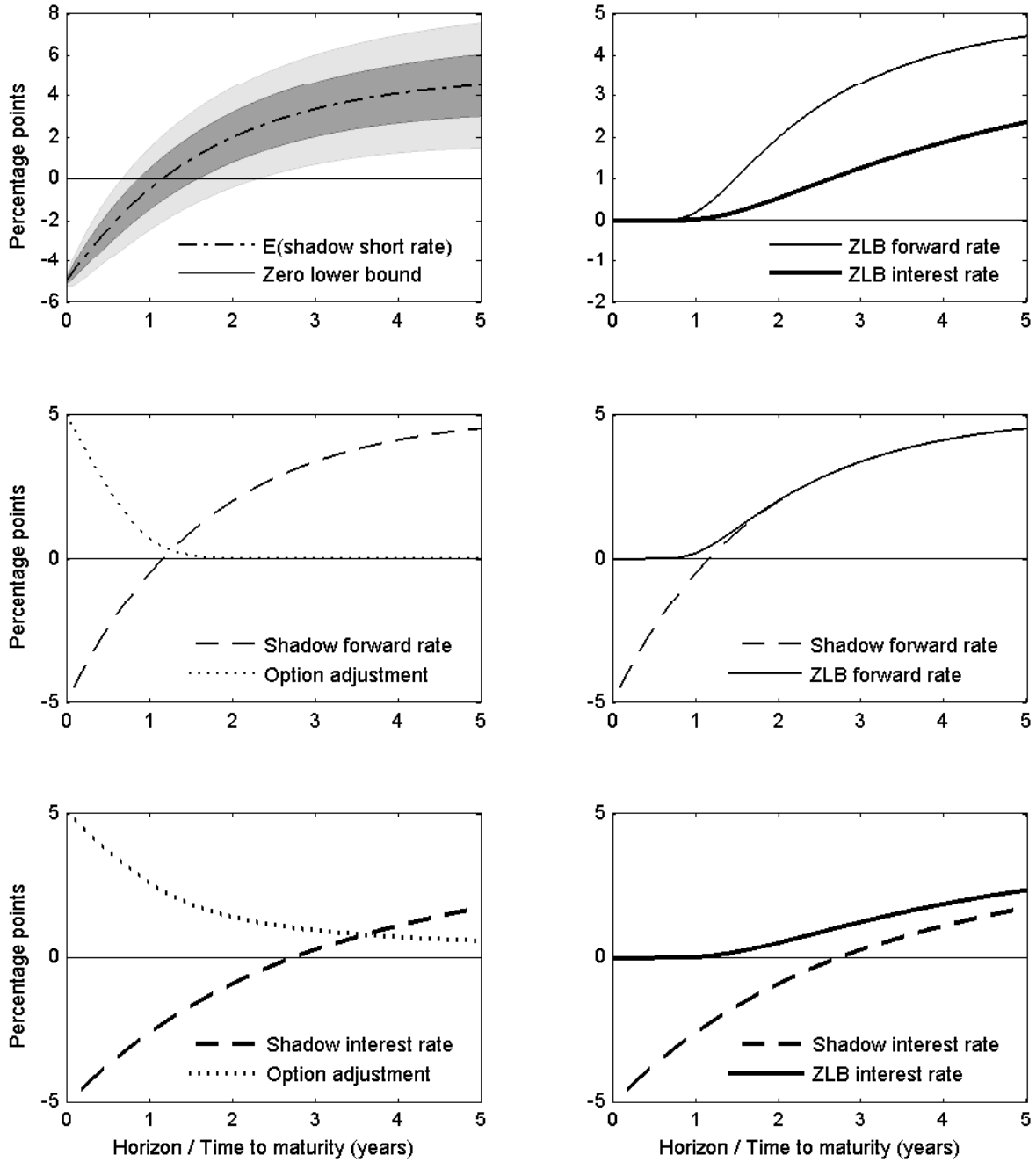


Figure 1: Perspectives on the ZLB-HJM term structure with $s_1(t) = 0.05$ and $s_1(t) = -0.10$, so $r(t) = -5$ percent.

As indicated in the top-left subplot, the expected path of the shadow short rate, as at time t and as a function of horizon τ (which I hereafter denote as $\mathbb{E}_t[r(t + \tau)]$) rises from its initial value of -5 percent to five percent (the value of $s_1(t)$, which sets the long-horizon level). However, $\mathbb{E}_t[r(t + \tau)]$ remains negative for around 1.2 years in this example. The shaded areas are respectively one and two standard deviation regions for the Gaussian probability densities that represent, as at time t , the expected distribution of future values of the shadow short rate as a function of horizon τ . For example, the probabilities of the shadow short rate being negative in around 1.5 years and 2.2 years time, respectively, are approximately $N[-1] = 0.159$ and $N[-2] = 0.023$.

The middle-left subplot shows the shadow forward rate $f(t, \tau)$ and the option-adjustment term $z(t, \tau)$. $f(t, \tau)$ is almost identical to $\mathbb{E}_t[r(t + \tau)]$ because the volatility-adjustment term $-\sigma_1^2 \cdot \frac{1}{2}\tau^2 - \sigma_2^2 \cdot \frac{1}{2}[G(\lambda, \tau)]^2$ is extremely small for short horizons.³⁰ Regarding $z(t, \tau)$, there are several points to note: (1) $z(t, \tau)$ takes on its highest value for $\tau = 0$, which in turn reflects the certainty that $r(t)$ is negative, so the option to hold physical currency is “in the money” and exercised (thereby offsetting the negative prevailing return of $r(t) = f(t, 0)$ that the investor faces on the shadow term structure); (2) $z(t, \tau)$ falls but remains materially positive out to around 1.5 years, which in turn reflects the high probabilities of $r(t)$ remaining negative over that horizon; and (3) $z(t, \tau)$ falls to and remains near zero as the expected path of the short rate becomes positive, reflecting the low probability of $r(t)$ remaining negative over moderate horizons.

The bottom-left subplot shows the shadow interest rate $r(t, \tau)$ and the interest rate option-adjustment component $\frac{1}{\tau} \int_0^\tau z(t, \tau) dv$. Being the cumulative averages of their forward rate counterparts, $r(t, \tau)$ and $\frac{1}{\tau} \int_0^\tau z(t, \tau) dv$ show similar patterns to $f(t, \tau)$ and $z(t, \tau)$, but they evolve more gradually by horizon/time to maturity τ .

The middle-right subplot shows the shadow forward rate $f(t, \tau)$ along with the associated ZLB forward rate $\underline{f}(t, \tau)$. The important points to note are: (1) $\underline{f}(t, \tau)$ remains essentially at zero until around 0.8 years (and it never takes on negative values); and (2) $\underline{f}(t, \tau)$ remains distinctly different from $f(t, \tau)$ to around 1.5 years. The bottom-right subplot shows shadow interest rates $r(t, \tau)$ and ZLB interest rates $\underline{r}(t, \tau)$, which again have similar but more gradual profiles relative to their forward rate counterparts. In particular, $\underline{r}(t, \tau)$ remains materially lower than $r(t, \tau)$ for the entire time-to-maturity span plotted.

Figure 2 shows that the ZLB-HJM term structure associated with a mildly negative shadow short rate is very similar to the shadow term structure itself. In this example, I set the ZLB-HJM state variables to $s_1(t) = 0.05$ and $s_2(t) = -0.055$, which results in a value for the current shadow short rate of $r(t) = s_1(t) + s_2(t) = -0.005$, or -0.5 percent. The expected path of the shadow short rate $\mathbb{E}_t[r(t + \tau)]$ again rises to five percent. In contrast to the example from figure 1, the ZLB-HJM forward rates and interest rates are immaterially different from their shadow counterparts. That, in turn, reflects the low probabilities of $r(t)$ remaining negative over short and moderate horizons. Note that an example with $r(t) = s_1(t) + s_1(t) = 0$ had ZLB-HJM forward rates and interest rates almost identical to the shadow counterparts.

Figure 3 shows that the ZLB-GATSM framework also accommodates negative forward rates that can arise for very long maturities in GATSMs with very persistent (i.e. slowly mean-reverting) state variables. The HJM model contains a unit root process for the Level state variable $s_1(t)$, which was obtained as the limit of zero mean reversion, or $\kappa_1 \rightarrow 0$. For this example, I set the ZLB-HJM state variables to $s_1(t) = 0.05$ and

³⁰I use the expression “volatility-adjustment term” because $\sigma_1^2 \cdot \frac{1}{2}\tau^2 + \sigma_2^2 \cdot \frac{1}{2}[G(\lambda, \tau)]^2$ is the difference between $f(t, \tau)$ and $\mathbb{E}_t[r(t + \tau)]$, and it represents the influence of volatility on expected returns due to Jensen’s inequality. That is, the expected compounded return from investing in a volatile short rate is less than the compounded return from investing in the expected (or mean) short rate.

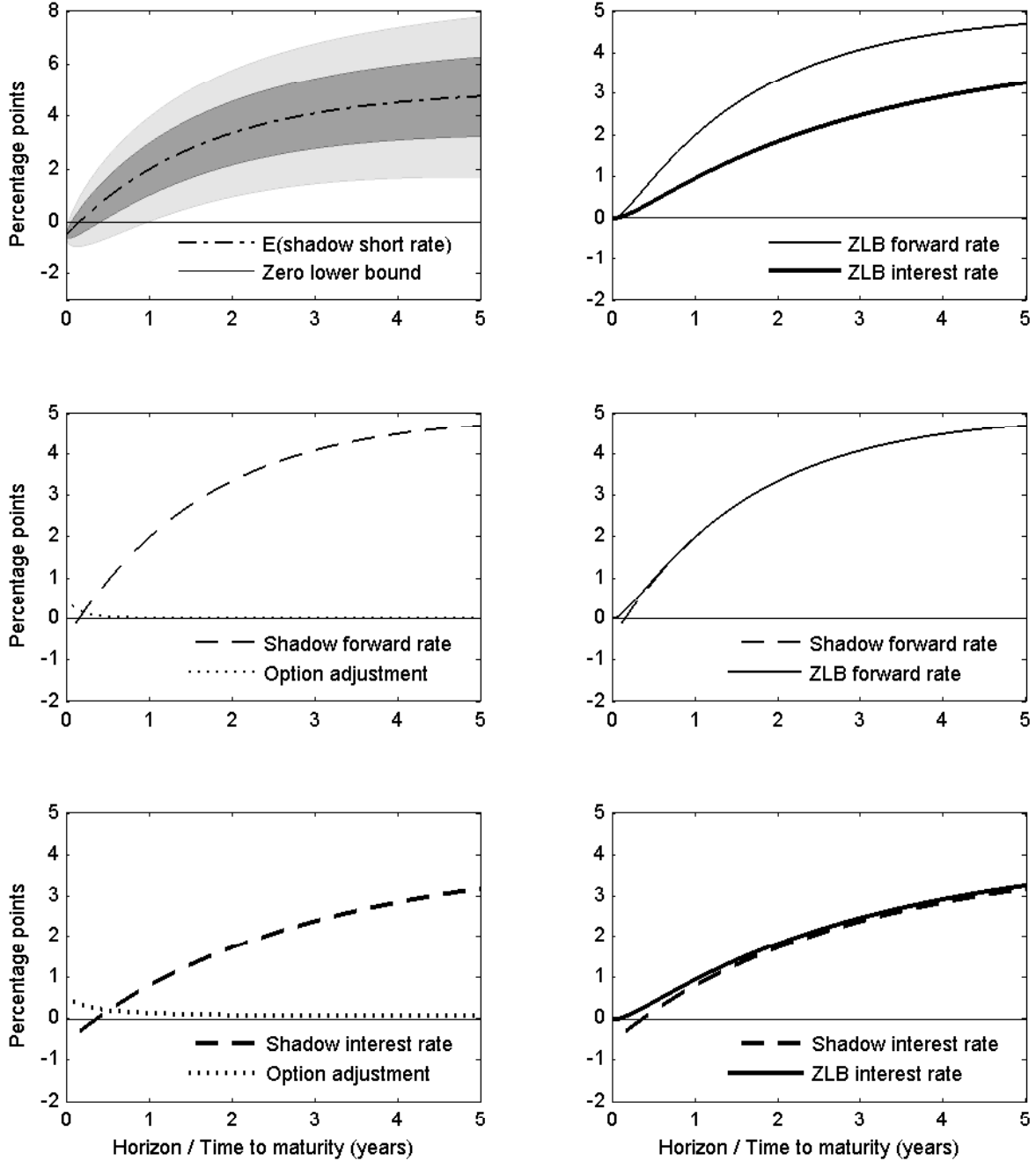


Figure 2: Perspectives on the ZLB-HJM term structure with $s_1(t) = 0.05$ and $s_2(t) = -0.055$, so $r(t) = -0.5$ percent.

$s_2(t) = 0$, which results in a value for the current shadow short rate of $r(t) = s_1(t) + s_2(t) = 0.05$, or five percent. With $s_2(t) = 0$, the expected path of the short rate $\mathbb{E}_t[r(t + \tau)]$ is also five percent for all horizons (although the results would be insensitive to any other reasonable value for $s_2(t)$, because the Slope factor loading effectively becomes zero for the long horizons being considered in this illustration).

For long times to maturity τ , it is apparent that the shadow forward rate $f(t, \tau)$ falls materially below $\mathbb{E}_t[r(t + \tau)]$, which is due to the increasing effect of the quadratic time-to-maturity term $-\sigma_1^2 \cdot \frac{1}{2}\tau^2$ in the volatility-adjustment term $-\sigma_1^2 \cdot \frac{1}{2}\tau^2 - \sigma_2^2 \cdot \frac{1}{2}[G(\lambda, \tau)]^2$.

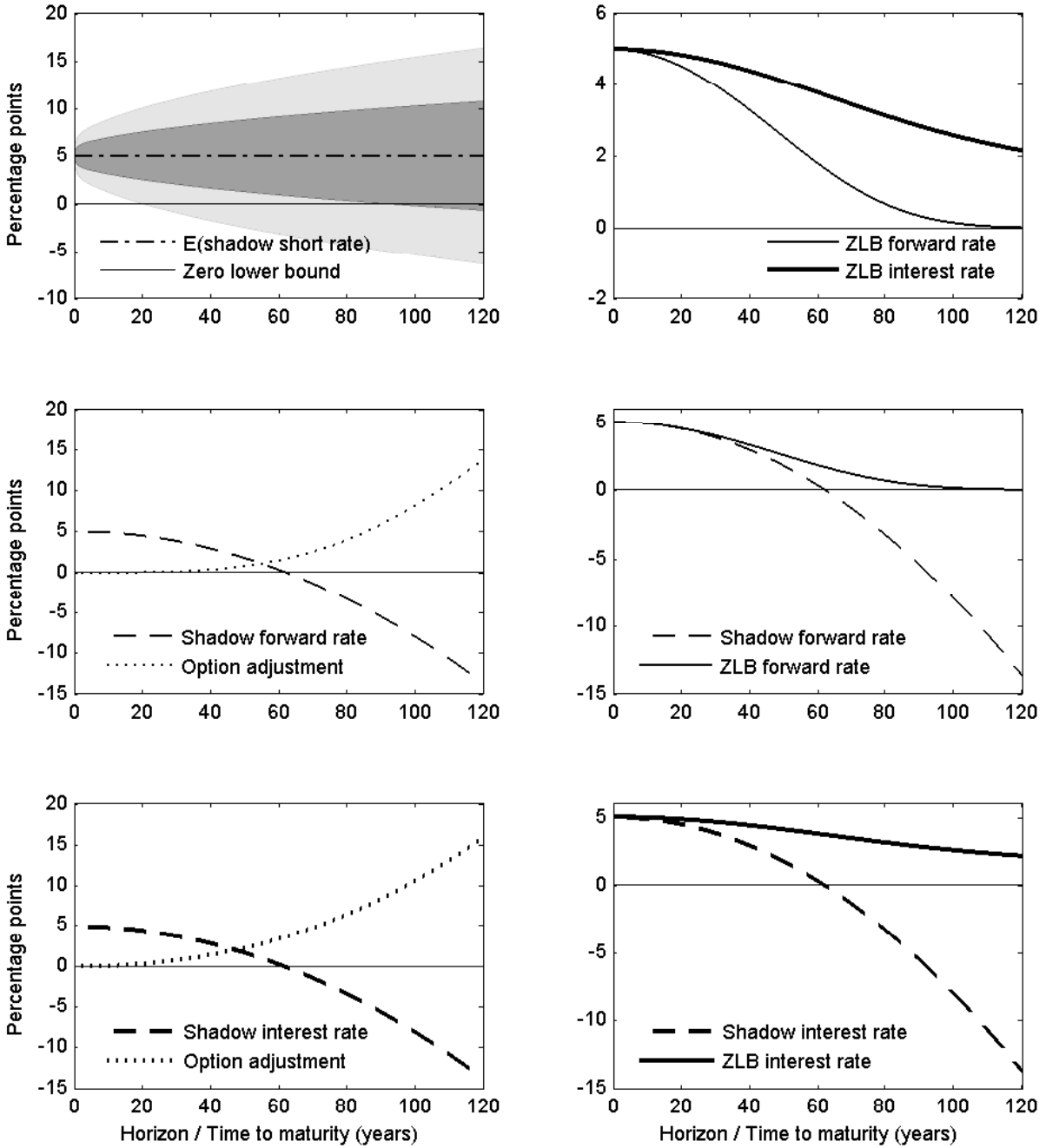


Figure 3: Perspectives on the ZLB-HJM term structure with $s_1(t) = 0.05$ and $s_1(t) = 0$, so $r(t)$ and $\mathbb{E}_t[r(t + \tau)] = 5$ percent.

For very long times to maturity (beyond 60 years in this example), it is a mathematical inevitability that the volatility-adjustment term will become large enough to result in negative forward rates (and, being a cumulative average, eventually negative interest rates). The option-adjustment term in the ZLB-GATSM framework resolves the issue by causing ZLB forward rates (and so ZLB interest rates) to asymptote to zero. Of course, whether that represents plausible values for very-long-horizon forward rates is an open question, but at least the process remains theoretically consistent with the presence of a ZLB at all horizons/times to maturity.

4.3 Dynamic/time-series perspective

This section demonstrates the time series perspective of the ZLB-GATSM framework by applying the ZLB-HJM model specified in section 4.1 to an illustrative estimation with U.S. term structure data. The objective is to show the evolution of the relative levels and dynamics of the ZLB-HJM state variables and the associated implied shadow short rate. The outright levels of those quantities should not be interpreted literally because, as will be discussed in section 5, the calibrated risk-neutral ZLB-HJM model used for this exercise is an over-parsimonious representation of term structure data observed in practice. In particular, the magnitudes of the negative implied shadow short rates will be overstated.

The data I use are end of month 3- and 6-month U.S. Treasury bill rates (from the Federal Reserve Economic Database on the St. Louis Federal Reserve website, converted to a continuously compounding basis) and 1-, 2-, 3-, 4-, 5-, 7-, 10-, 15-, and 30-year continuously compounding zero-coupon U.S. Treasury bond rates from the data set described in Gürkaynak, Sack, and Wright (2008). The sample starts in January 1987, simply to match the beginning of the CDR sample, and finishes at September 2011 (the last observation at the time of writing).

With the parameters λ , σ_1 , and σ_2 calibrated, the estimation may simply be undertaken via non-linear least squares at each point in time. Specifically, the objective is to optimize the following system:

$$\begin{bmatrix} \underline{r}(t, \tau_1) \\ \vdots \\ \underline{r}(t, \tau_K) \end{bmatrix} = \begin{bmatrix} \underline{r}(\{s_1(t), s_2(t)\}, \lambda, \sigma_1, \sigma_2, \tau_1) \\ \vdots \\ \underline{r}(\{s_1(t), s_2(t)\}, \lambda, \sigma_1, \sigma_2, \tau_K) \end{bmatrix} + \begin{bmatrix} \varepsilon(t, \tau_1) \\ \vdots \\ \varepsilon(t, \tau_K) \end{bmatrix} \quad (47a)$$

$$\text{minimize} \quad : \quad \text{SSR} = \sum_{k=1}^K [\varepsilon(t, \tau_k)]^2 \quad (47b)$$

where $\underline{r}(t, \tau_k)$ are the zero-coupon interest rate data at time t for time to maturity τ_k , $\underline{r}(\{s_1(t), s_2(t)\}, \lambda, \sigma_1, \sigma_2, \tau_k)$ are the estimated ZLB-HJM interest rates as a function of the parameters λ , σ_1 , σ_2 , and τ_k and the estimated state variables $s_1(t)$ and $s_2(t)$ at time t , $\varepsilon(t, \tau_k)$ are the differences between $\underline{r}(t, \tau_k)$ and $\underline{r}(\{s_1(t), s_2(t)\}, \lambda, \sigma_1, \sigma_2, \tau_k)$, and k runs from 1 to K indexing each data point being used to represent the term structure at time t . As an example, figure 4 illustrates the results from estimating the state variables for the last observation of term structure data in the sample, i.e. September 2011.

Figure 5 shows the 30-year interest rate data and the estimated time series for the ZLB-HJM Level state variable $s_1(t)$. These match closely over the entire sample, which is not surprising because the 30-year interest rate and the ZLB-HJM Level state variable $s_1(t)$ both reflect the long-run level of the term structure.³¹

Figure 6 shows the 3-month less 30-year interest rate spread and the estimated time series for the ZLB-HJM Slope state variable $s_2(t)$. These match fairly closely over the sample, except around 2003 and from late 2008. I discuss those deviations below from the perspectives of figures 7 and 8.

³¹Diebold and Li (2006) contains further discussion on the intuition of the Level state variable from the perspective of a non-arbitrage-free Nelson and Siegel (1987) model.

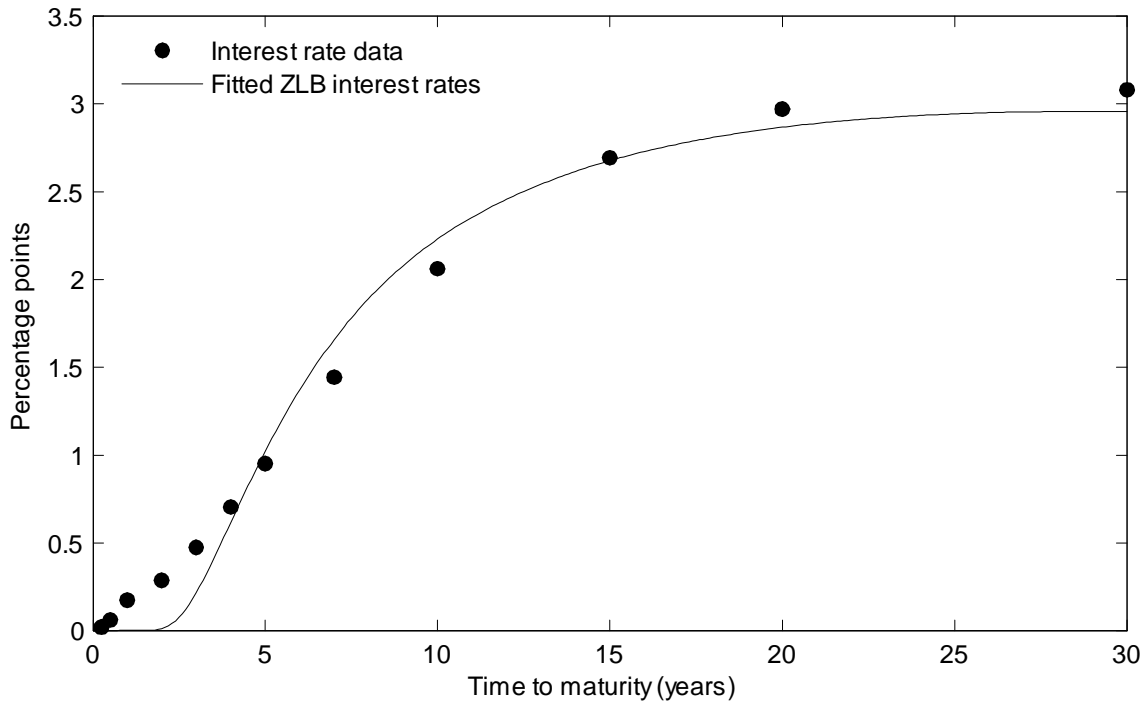


Figure 4: The Level and Slope state variables are respectively $s_1(t) = 0.0377$ and $s_2(t) = -0.1597$.

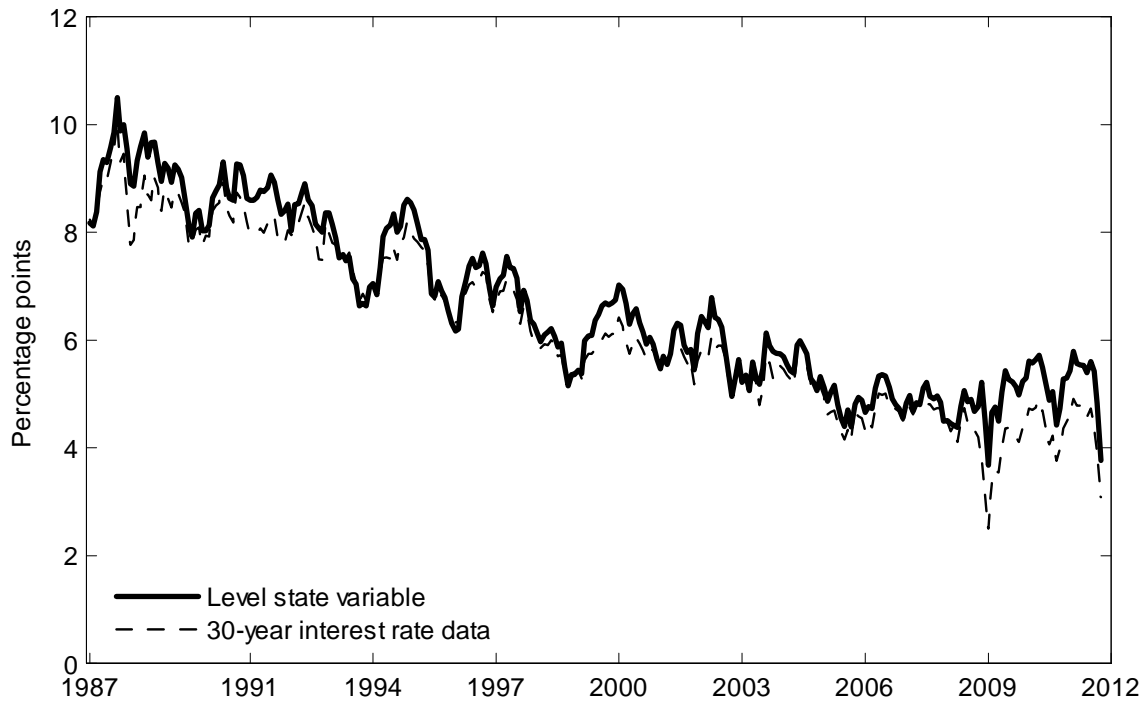


Figure 5: The 30-year interest rate and estimates of the the Level state variable $s_1(t)$ for the ZLB-HJM model.

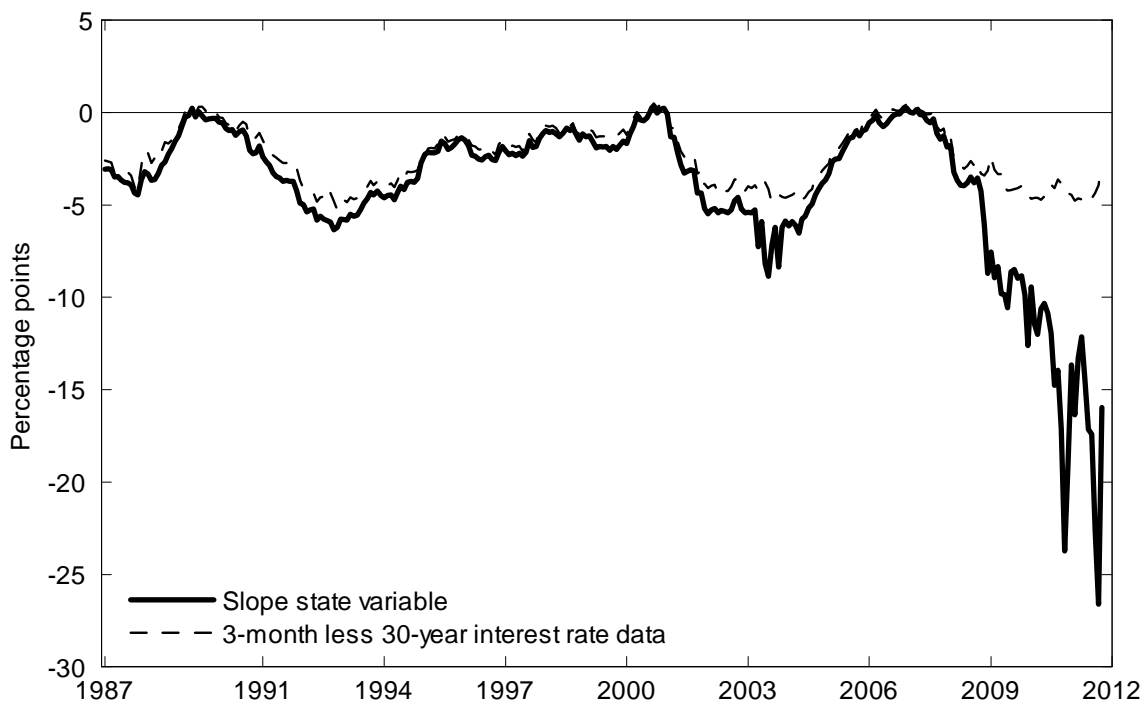


Figure 6: The 3-month less 30-year interest rate spread and estimates of the the Slope state variable $s_2(t)$ for the ZLB-HJM model.

Figures 7 and 8 show the 3-month interest rate data and the estimated time series for the shadow short rate $r(t) = s_1(t) + s_2(t)$.

The first point to note from figure 7 is that the estimated shadow short rate $r(t)$ is usually similar to the 3-month interest rate, as would be expected when the term structure data is not constrained by the ZLB. Second, $r(t)$ can freely adopt negative values (as allowed for within the ZLB-GATSM framework) while the 3-month interest rate can only drop to the ZLB, whereupon it is constrained from falling further (consistent with the “real world” availability of physical currency as an alternative investment). Third, the first instance of negative values for $r(t)$ occurred during the easing cycle that started with a 0.5 percentage point cut on 3 January 2001, and the most negative value for that cycle was $r(t) = -3.4$ percent in June 2003. That month coincided with the Federal Open Market Committee (hereafter FOMC) announcing a 0.25 percentage point cut to the federal funds rate (hereafter FFR) to 1 percent, which was the low point of that easing cycle, accompanied with a view that “*an unwelcome substantial fall in inflation exceeds that of a pickup in inflation from its already low level*” and “*the latter concern is likely to predominate for the foreseeable future.*”

Fourth, the most substantial period of negative values of $r(t)$ occurs from late 2008, which corresponds with events following the Global Financial Crisis that began in 2007. The shorter time span of figure 8 provides a better perspective on the downward evolution of the shadow short rate from 2007.

Figure 8 first indicates that the shadow short rate becomes negative in October 2008. Second, downward movements in the shadow short rate after $r(t)$ becomes negative are broadly consistent with events related to unconventional monetary policy easings that I have annotated in figure 7. For example, $r(t)$ drops substantially during November 2008, which coincides with the FOMC’s unscheduled announcement of “QE1” (the first program of “quantitative easing” [as labeled by the market] involving Federal Reserve

purchases of mortgage-backed agency securities and other measures). Similarly, $r(t)$ drops substantially again in the second half of 2010, following an FOMC warning about a slowing economic recovery and market anticipation of a second round of “QE2” (the second program of “quantitative easing”, involving Federal Reserve purchases of U.S. Treasury securities, subsequently announced in November 2010). The most recent substantial drop in $r(t)$, during 2011, follows recent FOMC warnings about a slowing economic recovery, the FOMC’s conditional commitment to keep the federal funds rate exceptionally low until mid-2013, and the (market-anticipated) FOMC announcement of a maturity extension program for the Federal Reserve’s balance sheet. Note that the sharp rise in $r(t)$ from late-2010 to early-2011 coincided with a run of encouraging economic data, which led markets to bring forward expectations of returning to a more “normal” economic environment and associated monetary policy settings. For example, over that period, the Standard and Poors 500 index appreciated by more than 20 percent and the 10-year bond yield rose more than one percentage point.

The third point from figure 8 is that movements in the shadow short rate appear to become more volatile after it becomes negative. This might be a genuine phenomenon, perhaps reflecting that a non-negative policy rate is more transparent to market participants and easier to target by the Federal Reserve than a negative shadow rate. Or the volatility increase might be an artifact of the estimated shadow short rate being influenced from the longer-maturity interest rates that tend to be more volatile.

In summary, the shadow short rate appears to move consistently with the FOMC’s stance of monetary policy and/or the market’s anticipation of that stance. That is, the shadow short rate closely tracks the 3-month rate when the term structure is not constrained by the ZLB, and the shadow short rate also reflects unconventional monetary policy easings, by taking on increasingly negative values, when the term structure becomes constrained by the ZLB.

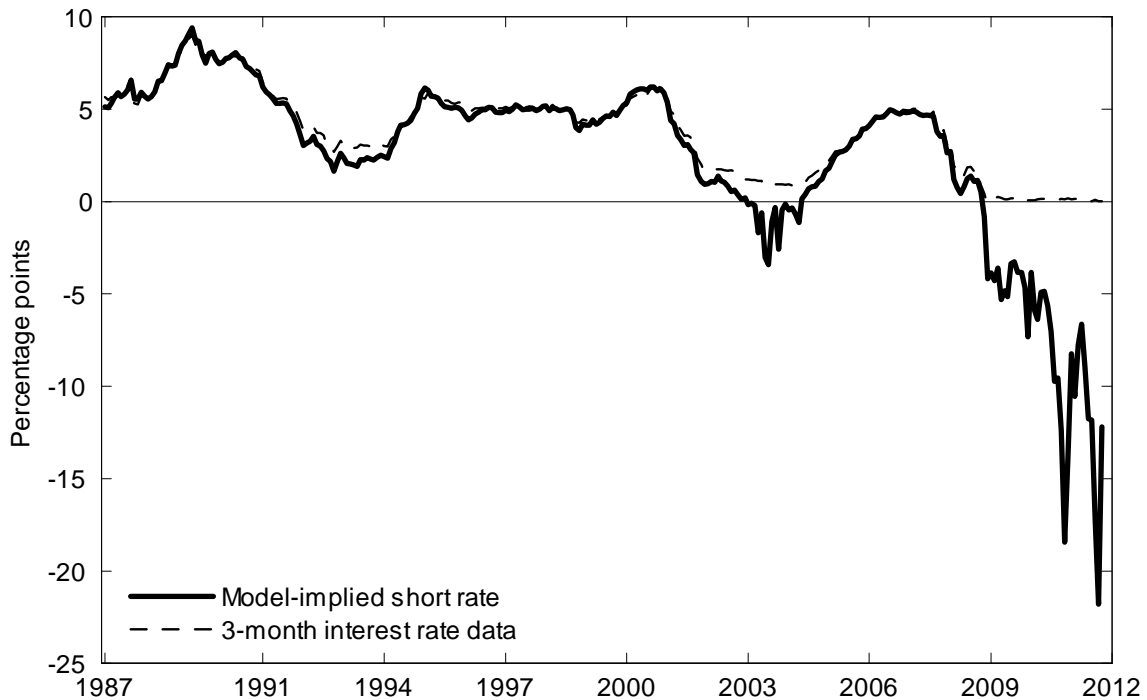


Figure 7: The 3-month interest rate and estimates of the short rate $r(t) = s_1(t) + s_2(t)$ for the ZLB-HJM model.

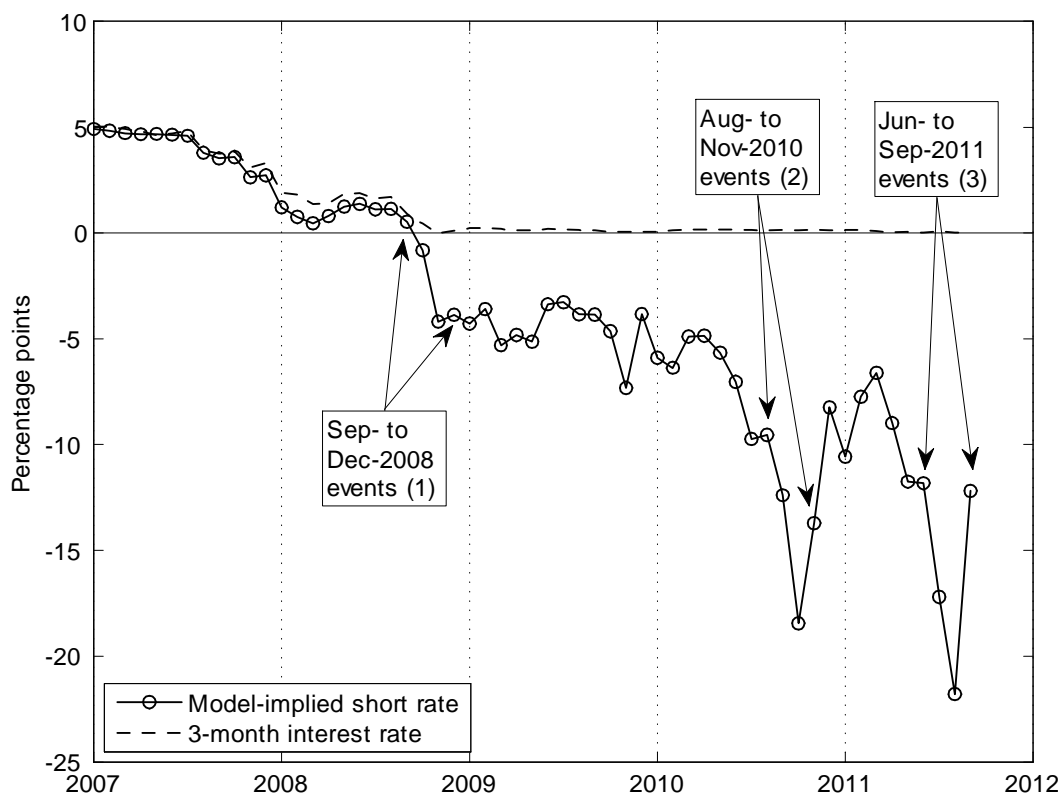


Figure 8: Estimates of the shadow short rate $r(t) = s_1(t) + s_2(t)$ for the ZLB-HJM model over the period January 2007 to September 2011. The annotated events associated with monetary policy easing are summarized below.

Events associated with unconventional monetary policy easings as indicated in figure 8:

- (1) 15-Sep: Lehman bankruptcy, 8-Oct: FFR cut from 2 to 1.5 percent, 29-Oct: FFR cut to 1 percent (unscheduled), 25-Nov: “QE1” announced (unscheduled), 16-Dec: FFR cut to 0-0.25 percent range.
- (2) 10-Aug: FOMC: “*recovery ... has slowed*”, 27-Aug: FOMC Chairman Bernanke, at Jackson Hole, foreshadows “QE2”, 3-Nov: QE2 announced.
- (3) 22-Jun: “*recovery is continuing ... though somewhat more slowly than the Committee had expected.*”, 9-Aug: “*The Committee currently anticipates ... exceptionally low levels for the federal funds rate at least through mid-2013*”, 26-Aug: FOMC Chairman Bernanke, at Jackson Hole, announces the September FOMC meeting will be two days “*instead of one to allow a fuller discussion*” of “*a range of tools that could be used to provide additional monetary stimulus*”. 21-Sep: FOMC announces a maturity extension program for the Federal Reserve’s balance sheet to “*put downward pressure on longer-term interest rates and help make broader financial conditions more accommodative.*”

5 Extensions and future work

The primary extension of the illustrative estimation in this article that I intend to undertake is a complete empirical estimation of the state variables and model parameters of the ZLB-HJM model. However, the HJM specification for the shadow term structure requires at least the addition of a process for the market prices of risk. The addition is required for the model to match the interest rate data which are observed under the physical or \mathbb{P} measure (where the market prices of risk are non-zero, so allowing for risk premiums), rather than the risk-neutral or \mathbb{Q} measure (with market prices of risk set to zero, so with no allowance for risk premiums). Note that the absence of risk premiums in the ZLB-HJM model used in this article leads to the magnitudes of the estimated negative shadow short rates being overstated. Essentially, the Slope state variable $s_2(t)$ is being “forced” to take on more negative values to account for term premiums in the data, thereby resulting in more negative values for the estimated shadow short rate. Another useful addition to the HJM specification would be an allowance for non-zero innovation correlations.

Others aspects to consider when seeking the “best” ZLB-GATSM for given applications would be the number of factors, whether one factor should be restricted to have a mean reversion rate of zero (as in the ZLB-HJM model), and whether time-invariant or time-varying market prices of risk are most appropriate.³² Those aspects are usually determined by assessing the fit to the term structure data and/or forecasting exercises.

Beyond specification, the estimation of the ZLB-GATSM model parameters and state variables would require a technique that allows for the term structure data being a non-linear function of the state variables. One of the variety of non-linear Kalman filters would likely prove suitable. It would also be desirable to allow the estimated parameters to vary over time, or at least undertake subsample estimation to allow for regime changes suggested by external information. For example, the apparent increase in volatility of the shadow short rate in figure 7 after the FFR reaches the ZLB might represent a testable regime change.

Another aspect related to estimation is the numerical integration method used to obtain ZLB-GATSM interest rates. I have used the most elementary method, but other numerical integration methods may prove more suitable. In addition, it would be straightforward to develop a hybrid GATSM/ZLB-GATSM framework that would use analytic solutions for interest rates and bond prices when the state variables are sufficiently different from zero (based on a materiality threshold, such as the 5/10000th of a basis point chosen in my examples), but revert to the ZLB-GATSM otherwise.

Potential extensions of the ZLB-GATSM framework would be to relax some of the simplifying assumptions if particular applications might benefit from that flexibility. For example, if investors ascribe some overhead cost to holding physical currency as a financial asset (e.g. due to risks of theft or insurance/protection expenses), that could readily be allowed for by setting the strike price of the options to a value above 1 (thereby allowing for slightly negative interest rates). And allowances could be made for different rates of borrowing and lending (e.g. to represent central bank floors and ceilings for overnight cash rates) by using separate borrowing and lending term

³²The BE/CIR economy specifies constant market prices of risk, but Berardi (2009) allows the market prices of risk specification to be “essentially affine” (i.e. a linear combination of the state variables), as introduced by Duffee (2002).

structures.

Regarding ZLB frameworks in general, a potential extension would be to use other models to represent the shadow term structure. In particular, some ZLB affine term structure model frameworks could readily be developed analogous to the ZLB-GATSM framework already established in the present article, because: (1) the BE/CIR economy already provides the explicit general equilibrium foundation for the Duffie and Kan (1996) and Dai and Singleton (2000) classes of multifactor affine term structure models, which includes Gaussian, square-root and mixed Gaussian/square-root dynamics;³³ and (2) closed-form analytic expressions for forward rates and option prices are already available for term structure models with independent square-root processes. Therefore, closed-form analytic expressions for ZLB forward rates would result for a generic multifactor model with independent square-root processes or mixed models combining the generic GATSM with one or several independent square-root processes. The latter might conveniently provide for some time-varying volatility (i.e. heteroskedasticity) while retaining most of the tractability of ZLB-GATSM framework. Alternatively, specifying a shadow term structure that allows for jump diffusions and/or stochastic volatility would also allow for heteroskedasticity, but potentially at the cost of some tractability.

6 Conclusion

In this article I have developed a generic framework for imposing a zero lower bound (ZLB) on Gaussian affine term structure models (GATSMs). Models within the ZLB-GATSM framework are theoretically consistent because they eliminate the probability of negative interest rates via an arbitrage-free shadow term structure that is founded on a general equilibrium economy. At the same time, ZLB-GATSMs retain all of the flexibility of GATSMs on which they are based, and ZLB-GATSMs retain most of the tractability of GATSMs. That is, ZLB forward rates have a simple closed-form analytic expression, and only elementary univariate numerical integration (over time to maturity) is required to obtain ZLB interest rates and ZLB bond prices.

I have shown with an illustrative application of a two-factor ZLB-GATSM to U.S. term structure data that movements in the model state variables have been consistent with unconventional monetary policy easings undertaken after the U.S. policy rate reached the ZLB in late 2008. However, a complete estimation with a richer specification remains as an important empirical extension.

References

- Berardi, A. (2009). Term structure, inflation, and real activity. *Journal of Financial and Quantitative Analysis* 44(4), 987–1011.
- Berardi, A. and M. Esposito (1999). A base model for multifactor specifications of the term structure. *Economic Notes* 28(2), 145–170.
- Black, F. (1995). Interest rates as options. *Journal of Finance* 50, 1371–1376.

³³I have simply used the purely Gaussian BE/CIR specification for my ZLB-GATSM framework.

- Chaplin, G. (1987). The term structure of interest rates - a model based on two correlated stochastic processes with closed form solutions for bond and option prices. *Working Paper, Department of Statistics and Actuarial Science, University of Waterloo ACTSC 87-16*.
- Chen, R. (1995). A two-factor, preference-free model for interest rate sensitive claims. *Journal of Futures Markets 15(3)*, 345–372.
- Christensen, J., F. Diebold, and G. Rudebusch (2009). An arbitrage-free generalized Nelson-Siegel term structure model. *Econometrics Journal 12(3)*, 33–64.
- Christensen, J., F. Diebold, and G. Rudebusch (2011). The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics 64*, 4–20.
- Cox, J., J. Ingersoll, and S. Ross (1985a). An intertemporal general equilibrium model of asset prices. *Econometrica 53*, 363–384.
- Cox, J., J. Ingersoll, and S. Ross (1985b). A theory of the term structure of interest rates. *Econometrica 53*, 385–407.
- Dai, Q. and K. Singleton (2000). Specification analysis of affine term structure models. *Journal of Finance 55(5)*, 1943–1978.
- Dai, Q. and K. Singleton (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics 63*, 415–441.
- Diebold, F. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130(2)*, 337–364.
- Duffee, G. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance 57(1)*, 405–443.
- Duffie, D. (1996). *Dynamic Asset Pricing Theory, Third Edition*. Princeton University Press.
- Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance 6(4)*, 379–406.
- Filipović, D. (2009). *Term-Structure Models: A Graduate Course*. Springer.
- Gorovoi, V. and V. Linetsky (2004). Black’s model of interest rates as options, eigenfunction expansions and Japanese interest rates. *Mathematical Finance 14(1)*, 49–78.
- Gürkaynak, R., B. Sack, and J. Wright (2008). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics 54*, 2291–2304.
- Hamilton, J. and J. Wu (2010). Identification and estimation of affine-term-structure models. *Working paper, University of California, San Diego*.
- Haug, E. (2007). *The Complete Guide to Option Pricing Formulas*. McGraw Hill.
- Heath, D., R. Jarrow, and A. Morton (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica 60(1)*, 77–106.
- Ichiue, H. and Y. Ueno (2006). Monetary policy and the yield curve at zero interest: the macro-finance model of interest rates as options. *Working paper, Bank of Japan 06-E-16*.

- Ichiue, H. and Y. Ueno (2007). Equilibrium interest rates and the yield curve in a low interest rate environment. *Working paper, Bank of Japan 07-E-18*.
- James, J. and N. Webber (2000). *Interest Rate Modelling*. Wiley and Sons.
- Joslin, S., K. Singleton, and H. Zhu (2010). A new perspective on Gaussian dynamic term structure models. *Review of Financial Studies (forthcoming)*.
- Kim, D. and K. Singleton (2011). Term structure models and the zero bound: an empirical investigation of Japanese yields. *Working paper, Stanford University*.
- Langetieg, T. (1980). A multivariate model of the term structure. *Journal of Finance 35(1)*, 71–97.
- Nelson, C. and A. Siegel (1987). Parsimonious modelling of yield curves. *Journal of Business October*, 473–489.
- Nelson, S. (1904). *The A B C of Options and Arbitrage*. New York: The Wall Street Library.
- Piazzesi, M. (2010). Affine term structure models. In L. Hansen and Y. A it-Sahalia (Eds.), *Macroeconomics, Finance and Econometrics: Essays in Memory of Albert Ando*, pp. 691–766. Elsevier.
- Rudebusch, G. (2010). Macro-finance models of interest rates and the economy. *The Manchester School Supplement 2010*, 25–52.
- Sharp, K. (1987). Bond and bond option prices dependent on n correlated stochastic processes. *Working Paper, Department of Statistics and Actuarial Science, University of Waterloo ACTSC 87-17*.
- Singleton, K. (2006). *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. Princeton University Press.
- Ueno, Y., N. Baba, and Y. Sakurai (2006). The use of the Black model of interest rates as options for monitoring the JGB market expectations. *Working paper, Bank of Japan 06-E-15*.
- Vasicek, O. (1977). An equilibrium characterisation of the term structure. *Journal of Financial Economics 5*, 177–188.
- Vincente, J. and B. Tabak (2008). Forecasting bond yields in the Brazilian fixed income market. *International Journal of Forecasting 24*, 490–497.

A ZLB forward rate derivation

This appendix details the derivation of the ZLB forward rate expression in section 2.3.2.

$$\begin{aligned}
\underline{f}(t, \tau) &= \lim_{\delta \rightarrow 0} \underline{F}(t, \tau, \tau + \delta) \\
&= \lim_{\delta \rightarrow 0} \left\{ -\frac{1}{\delta} \log \left[\frac{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right\} \\
\langle \text{L'Hopital's rule} \rangle &= \lim_{\delta \rightarrow 0} \left\{ -\frac{d}{d\delta} \left(\log \left[\frac{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right) \right\} \\
\langle \text{Chain rule} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{dx} \log [x] \cdot -\frac{d}{d\delta} \left[\frac{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right\} \\
&\quad \left\langle x = \frac{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)}{P(t, \tau)}; \frac{d}{dx} \log [x] = \frac{1}{x} \right\rangle \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau)}{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)} \right. \\
&\quad \left. \times \left(-\frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] + \frac{d}{d\delta} \left[\frac{C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right) \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau)}{P(t, \tau + \delta) - C(t, \tau, \tau + \delta)} \right\} \\
&\quad \times \left[\lim_{\delta \rightarrow 0} \left\{ -\frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} + \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(\cdot)}{P(t, \tau)} \right] \right\} \right] \quad (48)
\end{aligned}$$

where ‘‘Limit rules’’ refers to the standard rules for manipulating limits (in this case their sums and products; i.e. respectively, $\lim_{x \rightarrow a} \{f(x) + g(x)\} = \lim_{x \rightarrow a} f(x) + \lim_{x \rightarrow a} g(x)$, and $\lim_{x \rightarrow a} \{f(x) \cdot g(x)\} = \lim_{x \rightarrow a} f(x) \cdot \lim_{x \rightarrow a} g(x)$).

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau)}{P(t, \tau + \delta) - C(\cdot)} \right\}$ in equation 48 may be evaluated with the quotient limit rule (i.e. $\lim_{x \rightarrow a} \{f(x) / g(x)\} = \lim_{x \rightarrow a} f(x) / \lim_{x \rightarrow a} g(x)$ if $\lim_{x \rightarrow a} g(x) \neq 0$) as follows:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau)}{P(t, \tau + \delta) - C(\cdot)} \right\} &= \frac{\lim_{\delta \rightarrow 0} \{P(t, \tau)\}}{\lim_{\delta \rightarrow 0} \{P(t, \tau + \delta) - C(\cdot)\}} \\
\langle \text{Limit rules} \rangle &= \frac{P(t, \tau)}{\lim_{\delta \rightarrow 0} \{P(t, \tau + \delta)\} - \lim_{\delta \rightarrow 0} \{C(\cdot)\}} \\
&= \frac{P(t, \tau)}{P(t, \tau) - 0} \\
&= 1 \quad (49)
\end{aligned}$$

where the result $\lim_{\delta \rightarrow 0} \{C(t, \tau, \tau + \delta)\} = 0$ follows from taking the limit for the generic expression $C(t, \tau, \tau + \delta)$ in 13, and noting that

$$\begin{aligned}
\max \left[\lim_{\delta \rightarrow 0} P(t + \tau, \delta) - 1, 0 \right] &= \max [P(t + \tau, 0) - 1, 0] \\
&= \max [1 - 1, 0] \\
&= 0 \quad (50)
\end{aligned}$$

The expression $\lim_{\delta \rightarrow 0} \left\{ -\frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\}$ in equation 48 may be evaluated by first calculating the following derivative:

$$\begin{aligned}
f(t, \tau + \delta) &= -\frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \\
\langle \text{Chain rule} \rangle &= -\frac{d}{dx} \log [x] \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \\
&\quad \left\langle x = \frac{P(t, \tau + \delta)}{P(t, \tau)}; \frac{d}{dx} \log [x] = \frac{1}{x} \right\rangle \\
&= -\frac{P(t, \tau)}{P(t, \tau + \delta)} \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right]
\end{aligned} \tag{51}$$

and then rearranging the result to give:

$$\frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] = -\frac{P(t, \tau + \delta)}{P(t, \tau)} \cdot f(t, \tau + \delta) \tag{52}$$

The evaluation of the limit is then:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} &= \lim_{\delta \rightarrow 0} \left\{ -\frac{P(t, \tau + \delta)}{P(t, \tau)} \cdot f(t, \tau + \delta) \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ -\frac{P(t, \tau + \delta)}{P(t, \tau)} \right\} \cdot \lim_{\delta \rightarrow 0} \{f(t, \tau + \delta)\} \\
&= -1 \cdot f(t, \tau) \\
&= -f(t, \tau)
\end{aligned} \tag{53}$$

Finally, I denote:

$$z(t, \tau) = \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right\} \tag{54}$$

where the result of this calculation will depend on the specific model used to represent the shadow term structure.

The final result for equation 48 is therefore:

$$\underline{f}(t, \tau) = f(t, \tau) + z(t, \tau) \tag{55}$$

B ZLB-GATSM forward rate derivation

This appendix contains the derivations of the expressions in sections 3.2.2 and 3.2.3.

B.1 Shadow/GATSM forward rates

GATSM forward rates may be obtained using the standard bond price/forward rate definition and equation 23:

$$\begin{aligned}
f(t, \tau) &= -\frac{d}{d\tau} \log [P(t, \tau)] \\
&= -\frac{d}{d\tau} \log \left[\exp \left(-H(\tau) - \sum_{n=1}^N s_n(t) \cdot G(\kappa_n, \tau) \right) \right] \\
&= \frac{d}{d\tau} H(\tau) + \sum_{n=1}^N s_n(t) \cdot \frac{d}{d\tau} G(\kappa_n, \tau)
\end{aligned} \tag{56}$$

The expressions $\frac{d}{d\tau} G(\kappa_n, \tau)$ are readily calculated from the definition in equation 24, i.e.:

$$\begin{aligned}
\frac{d}{d\tau} G(\kappa_n, \tau) &= \frac{d}{d\tau} \left(\frac{1}{\kappa_n} [1 - \exp(-\kappa_n \tau)] \right) \\
&= \frac{1}{\kappa_n} \cdot \kappa_n \exp(-\kappa_n \tau) \\
&= \exp(-\kappa_n \tau)
\end{aligned} \tag{57}$$

as is the expression $\frac{d}{d\delta} H(\tau)$ from the definition in equation 25, i.e.:

$$\begin{aligned}
\frac{d}{d\tau} H(\tau) &= \frac{d}{d\tau} \left(-\frac{1}{2} \text{Tr} [\tau \Xi(\tau) \Psi] + \sum_{n=1}^N \mu_n [\tau - G(\kappa_n, \tau)] \right) \\
&= -\frac{1}{2} \text{Tr} \left[\frac{d}{d\tau} [\tau \Xi(\tau) \Psi] \right] + \sum_{n=1}^N \mu_n \left[\frac{d}{d\tau} \tau - \frac{d}{d\tau} G(\kappa_n, \tau) \right]
\end{aligned} \tag{58}$$

$$= -\frac{1}{2} \text{Tr} [\Theta(\tau) \Psi] + \sum_{n=1}^N \mu_n [1 - \exp(-\kappa_n \tau)] \tag{59}$$

where:

$$\begin{aligned}
\Theta_{ij}(\tau) &= \frac{d}{d\tau} [\tau \Xi(\tau)]_{ij} \\
&= \rho_{ij} \sigma_i \sigma_j \cdot \frac{d}{d\tau} [\tau - G(\kappa_i, \tau) - G(\kappa_j, \tau) + G(\kappa_i + \kappa_j, \tau)] \\
&= \rho_{ij} \sigma_i \sigma_j \cdot [1 - \exp(-\kappa_i \tau) - \exp(-\kappa_j \tau) + \exp(-\{\kappa_i + \kappa_j\} \tau)] \\
&= \rho_{ij} \sigma_i \sigma_j \cdot [1 - \exp(-\kappa_i \tau)] [1 - \exp(-\kappa_j \tau)] \\
&= \rho_{ij} \sigma_i \sigma_j \cdot \kappa_i \kappa_j G(\kappa_i, \tau) G(\kappa_j, \tau)
\end{aligned} \tag{60}$$

The forward rate for the generic GATSM is therefore:

$$f(t, \tau) = -\frac{1}{2} \text{Tr} [\Theta(\tau) \Psi] + \sum_{n=1}^N \mu_n + \sum_{n=1}^N [s_n(t) - \mu_n] \exp(-\kappa_n \tau) \tag{61}$$

B.2 Option-adjustment term

The option-adjustment term may be obtained from its definition in equation 17 and the GATSM option price expressions in section 3.1.3. I first substitute the option price formula from equation 28 into equation 17 and then derive the resulting three limit expressions.

$$\begin{aligned}
z(t, \tau) &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \tau + \delta)}{P(t, \tau)} \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta) N[d_1(\cdot)] - P(t, \tau) N[d_2(\cdot)]}{P(t, \tau)} \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} N[d_1(\cdot)] - N[d_2(\cdot)] \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \cdot N[d_1(\cdot)] \right. \\
&\quad \left. + \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} N[d_1(\cdot)] - \frac{d}{d\delta} N[d_2(\cdot)] \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} \cdot \lim_{\delta \rightarrow 0} \{N[d_1(t, \tau, \tau + \delta)]\} \\
&\quad + \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} N[d_1(\cdot)] - \frac{d}{d\delta} N[d_2(t, \tau, \tau + \delta)] \right\} \quad (62)
\end{aligned}$$

The result $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} = -f(t, \tau)$ is already available from equation 53. The details for deriving the other two limits required for equation 62 are contained in the following subsections, with the following results:

$$\lim_{\delta \rightarrow 0} \{N[d_1(t, \tau, \tau + \delta)]\} = 1 - N \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \quad (63)$$

$$\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} N[d_1(\cdot)] - \frac{d}{d\delta} N[d_2(\cdot)] \right\} = \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \quad (64)$$

The final result for $z(t, \tau)$ is therefore:

$$z(t, \tau) = -f(t, \tau) \cdot \left(1 - N \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \quad (65)$$

B.2.1 $\lim_{\delta \rightarrow 0} \{N[d_1(t, \tau, \tau + \delta)]\}$

The definition of the normal distribution $N[x]$ in terms of the error function $\text{erf}[x]$ is as follows:

$$N[x] = \frac{1}{2} + \frac{1}{2} \text{erf} \left[\frac{1}{\sqrt{2}} x \right] \quad (66)$$

and the definition of the error function is itself:

$$\text{erf}[y] = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n [y]^{2n+1}}{n! (2n+1)} \quad (67)$$

Substituting $y = \frac{1}{\sqrt{2}}x$ and $x = d_1(t, \tau, \tau + \delta)$ into the preceding expressions gives:

$$N[d_1(t, \tau, \tau + \delta)] = \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} d_1(t, \tau, \tau + \delta) \right]^{2n+1}}{n! (2n+1)} \quad (68)$$

and taking the limit gives:

$$\begin{aligned} \lim_{\delta \rightarrow 0} N[d_1(\cdot)] &= \frac{1}{2} + \frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} d_1(\cdot) \right]^{2n+1}}{n! (2n+1)} \right\} \\ \langle \text{Limit rules} \rangle &= \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \cdot \lim_{\delta \rightarrow 0} \left\{ \left[\frac{1}{\sqrt{2}} d_1(\cdot) \right]^{2n+1} \right\}}{n! (2n+1)} \\ \langle \text{Limit rules} \rangle &= \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} \cdot \lim_{\delta \rightarrow 0} d_1(\cdot) \right]^{2n+1}}{n! (2n+1)} \end{aligned} \quad (69)$$

where ‘‘Limit rules’’ in the final line includes the power limit rule $\lim_{x \rightarrow a} \{[f(x)]^n\} = \{[\lim_{x \rightarrow a} f(x)]\}^n$.

The expression for $\lim_{\delta \rightarrow 0} d_1(t, \tau, \tau + \delta)$ is derived directly from the expression in equation 29, i.e:

$$\begin{aligned} \lim_{\delta \rightarrow 0} d_1(t, \tau, \tau + \delta) &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\Sigma(\tau, \tau + \delta)} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] + \frac{1}{2} \Sigma(\tau, \tau + \delta) \right\} \\ \langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\Sigma(\tau, \tau + \delta)} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} + \frac{1}{2} \lim_{\delta \rightarrow 0} \Sigma(\tau, \tau + \delta) \\ \langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\}}{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\}} + \frac{1}{2} \lim_{\delta \rightarrow 0} \Sigma(\tau, \tau + \delta) \end{aligned} \quad (70)$$

The result for $\lim_{\delta \rightarrow 0} \Sigma(\tau, \tau + \delta)$ is readily calculated as zero, by referring to equation 31 and noting that for any x :

$$\begin{aligned} \lim_{\delta \rightarrow 0} G(x, \delta) &= \lim_{\delta \rightarrow 0} \left\{ \frac{1 - \exp(-x\delta)}{x} \right\} \\ &= \frac{1 - 1}{x} \\ &= 0 \end{aligned} \quad (71)$$

Therefore:

$$\lim_{\delta \rightarrow 0} \Sigma(\tau, \tau + \delta) = 0 \quad (72)$$

Note that equation 72 also justifies the application of L'Hopital's rule to the first expression on the second-last line of equation 70, because the limit would be undefined (i.e. $\log [P(t, \tau) / P(t, \tau)] = \log [1] = 0$ and $\Sigma(\tau, \tau) = 0$).

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\}$ required for equation 70 is:

$$\begin{aligned} \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log [P(t, \tau + \delta)] - \frac{d}{d\delta} \log [P(t, \tau)] \right\} \\ &= \lim_{\delta \rightarrow 0} \{-f(t, \tau + \delta) - 0\} \\ &= -f(t, \tau) \end{aligned} \quad (73)$$

and $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\}$ is the expression for annualized instantaneous volatility, which I derive in section B.3 and denote $\omega(\tau)$, i.e.:

$$\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} = \omega(\tau) \quad (74)$$

Substituting these results into equation 70 gives:

$$\lim_{\delta \rightarrow 0} d_1(t, \tau, \tau + \delta) = -\frac{f(t, \tau)}{\omega(\tau)} \quad (75)$$

and the final result for $\lim_{\delta \rightarrow 0} N[d_1(t, \tau, \tau + \delta)]$ is obtained substituting that result into equation 69, i.e.:

$$\begin{aligned} \lim_{\delta \rightarrow 0} N[d_1(t, \tau, \tau + \delta)] &= \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} \cdot \left(-\frac{f(t, \tau)}{\omega(\tau)} \right) \right]^{2n+1}}{n! (2n+1)} \\ &= \frac{1}{2} + \frac{1}{2} \cdot \operatorname{erf} \left[-\frac{1}{\sqrt{2}} \frac{f(t, \tau)}{\omega(\tau)} \right] \\ &= N \left[-\frac{f(t, \tau)}{\omega(\tau)} \right] \\ &= 1 - N \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \end{aligned} \quad (76)$$

B.2.2 $\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} N[d_1(t, \tau, \tau + \delta)] - \frac{d}{d\delta} N[d_2(t, \tau, \tau + \delta)] \right\}$

I first derive the terms $\frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} N[d_1(t, \tau, \tau + \delta)]$ and $\frac{d}{d\delta} N[d_2(t, \tau, \tau + \delta)]$ separately, and then calculate the limit of their difference (which I later denote as $U(t, \tau, \tau + \delta)$ for notational convenience).

$$\begin{aligned} \frac{d}{d\delta} N[d_1(t, \tau, \tau + \delta)] &= \frac{d}{dx} N[x] \frac{d}{d\delta} d_1(t, \tau, \tau + \delta) \\ \langle \text{Chain rule} \rangle &: \left\langle x = d_1(t, \tau, \tau + \delta); \frac{d}{dx} N[x] = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} x^2 \right] \right\rangle \\ &= \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(t, \tau, \tau + \delta)]^2 \right] \frac{d}{d\delta} d_1(t, \tau, \tau + \delta) \end{aligned} \quad (77)$$

It turns out that $\frac{d}{d\delta} d_1(t, \tau, \tau + \delta)$ is not required in explicit form. Pre-multiplying by $\frac{P(t, \tau + \delta)}{P(t, \tau)}$ then gives:

$$\frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} N[d_1(t, \tau, \tau + \delta)] = \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(\cdot)]^2 \right] \frac{d}{d\delta} d_1(\cdot) \quad (78)$$

Analogous to the derivative for $N[d_1(t, \tau, \tau + \delta)]$:

$$\frac{d}{d\delta} N[d_2(t, \tau, \tau + \delta)] = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_2(t, \tau, \tau + \delta)]^2\right] \frac{d}{d\delta} d_2(t, \tau, \tau + \delta) \quad (79)$$

The term $[d_2(t, \tau, \tau + \delta)]^2$ may be re-expressed as follows:

$$\begin{aligned} [d_2(t, \tau, \tau + \delta)]^2 &= [d_1(t, \tau, \tau + \delta) - \Sigma(\tau, \tau + \delta)]^2 \\ &= [d_1(\cdot)]^2 - 2d_1(\cdot) \Sigma(\tau, \tau + \delta) + [\Sigma(\tau, \tau + \delta)]^2 \\ &= [d_1(\cdot)]^2 - 2 \left(\frac{1}{\Sigma(\cdot)} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] + \frac{1}{2} \Sigma(\cdot) \right) \Sigma(\cdot) - [\Sigma(\cdot)]^2 \\ &= [d_1(\cdot)]^2 - 2 \cdot \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] + [\Sigma(\cdot)]^2 - [\Sigma(\cdot)]^2 \\ &= [d_1(\cdot)]^2 - 2 \cdot \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \end{aligned} \quad (80)$$

and therefore:

$$\begin{aligned} \exp\left[-\frac{1}{2} [d_2(t, \tau, \tau + \delta)]^2\right] &= \exp\left[-\frac{1}{2} \left([d_1(\cdot)]^2 - 2 \cdot \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right)\right] \\ &= \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \cdot \exp\left[\log \left(\frac{P(t, \tau + \delta)}{P(t, \tau)} \right)\right] \\ &= \frac{P(t, \tau + \delta)}{P(t, \tau)} \exp\left[-\frac{1}{2} [d_1(t, \tau, \tau + \delta)]^2\right] \end{aligned} \quad (81)$$

The derivative $\frac{d}{d\delta} d_2(t, \tau, \tau + \delta)$ is:

$$\begin{aligned} \frac{d}{d\delta} d_2(t, \tau, \tau + \delta) &= \frac{d}{d\delta} [d_1(t, \tau, \tau + \delta) - \Sigma(\tau, \tau + \delta)] \\ &= \frac{d}{d\delta} d_1(t, \tau, \tau + \delta) - \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \end{aligned} \quad (82)$$

Substituting the results from equations 81 and 82 into equation 79 gives the following expression:

$$\begin{aligned} \frac{d}{d\delta} N[d_2(\cdot)] &= \frac{1}{\sqrt{2\pi}} \frac{P(t, \tau + \delta)}{P(t, \tau)} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \left[\frac{d}{d\delta} d_1(\cdot) - \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right] \\ &= \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta} d_1(\cdot) \\ &\quad - \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \end{aligned} \quad (83)$$

When calculating $U(t, \tau, \tau + \delta) = \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} N[d_1(t, \tau, \tau + \delta)] - \frac{d}{d\delta} N[d_2(t, \tau, \tau + \delta)]$, note that the expression $\frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta} d_1(\cdot)$ in equation 78 cancels identically with the first component from equation 83. Therefore:

$$U(t, \tau, \tau + \delta) = \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta} \Sigma(\cdot) \quad (84)$$

and taking the limit results in three more limit expressions to derive, i.e:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} U(t, \tau, \tau + \delta) &= \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(\cdot)]^2 \right] \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \right\} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(\cdot)]^2 \right] \right\} \\
&\quad \times \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} \tag{85}
\end{aligned}$$

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} = \omega(\tau)$, as derived and denoted in section B.3, and the expression $\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \right\}$ is readily derived, i.e:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \right\} &= \frac{P(t, \tau)}{P(t, \tau)} \\
&= 1 \tag{86}
\end{aligned}$$

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(\cdot)]^2 \right] \right\}$ is derived as follows. The definition of the exponential function is:

$$\exp[x] = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{87}$$

Substituting $x = -\frac{1}{2} [d_1(t, \tau, \tau + \delta)]^2$ into equation 87 and taking the limit gives:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(\cdot)]^2 \right] \right\} &= \frac{1}{\sqrt{2\pi}} \lim_{\delta \rightarrow 0} \left\{ \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2} [d_1(\cdot)]^2\right)^n}{n!} \right\} \\
&= \frac{1}{\sqrt{2\pi}} \lim_{\delta \rightarrow 0} \left\{ \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n [d_1(\cdot)]^{2n}}{n!} \right\} \\
\langle \text{Limit rules} \rangle &= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n \lim_{\delta \rightarrow 0} \{ [d_1(\cdot)]^{2n} \}}{n!} \\
\langle \text{Limit rules} \rangle &= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n [\lim_{\delta \rightarrow 0} d_1(t, \tau, \tau + \delta)]^{2n}}{n!} \tag{88}
\end{aligned}$$

The result $\lim_{\delta \rightarrow 0} [d_1(t, \tau, \tau + \delta)] = -\frac{f(t, \tau)}{\omega(\tau)}$ is available from equation 75, and so

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(t, \tau, \tau + \delta)]^2 \right] \right\} &= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n \left[-\frac{f(t, \tau)}{\omega(\tau)} \right]^{2n}}{n!} \\
&= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n \left(\left[-\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right)^n}{n!} \\
&= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right)^n}{n!} \\
&= \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \tag{89}
\end{aligned}$$

The final result for $\lim_{\delta \rightarrow 0} U(t, \tau, \tau + \delta)$ is therefore:

$$\begin{aligned} \lim_{\delta \rightarrow 0} U(t, \tau, \tau + \delta) &= 1 \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right) \cdot \omega(\tau) \\ &= \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right) \end{aligned} \quad (90)$$

B.3 Annualized instantaneous volatility

I define annualized instantaneous volatility $\omega(\tau)$ as the annualized limit of the option volatility expression from equation 31 and denote it $\omega(\tau)$, i.e.:

$$\begin{aligned} \omega(\tau) &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\delta} \Sigma(\tau, \tau + \delta) \right\} \\ \langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\}}{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \delta \right\}} \\ &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} \end{aligned} \quad (91)$$

because $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \delta \right\} = 1$.

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\}$ may be calculated as follows:

$$\begin{aligned} \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \sqrt{[\Sigma(\tau, \tau + \delta)]^2} \right\} \\ \langle \text{Chain rule} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{dx} \sqrt{x} \frac{d}{d\delta} [\Sigma(\tau, \tau + \delta)]^2 \right\} \\ &\quad \left\langle x = [\Sigma(\tau, \tau + \delta)]^2; \frac{d}{dx} \sqrt{x} = \frac{1}{2} x^{-1/2} = \frac{1}{2\sqrt{x}} \right\rangle \\ &= \lim_{\delta \rightarrow 0} \left\{ \frac{\frac{d}{d\delta} [\Sigma(\tau, \tau + \delta)]^2}{2\Sigma(\tau, \tau + \delta)} \right\} \\ \langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{d}{d\delta} [\Sigma(\tau, \tau + \delta)]^2 \right] \right\}}{2 \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\}} \\ &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \tau + \delta)]^2 \right\}}{2 \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\}} \end{aligned} \quad (92)$$

Note that L'Hopital's rule has been used in the second-last step of equation 92 because in the middle line both $\frac{d}{d\delta} [\Sigma(\tau, \tau + \delta)]^2 = 2\Sigma(\tau, \tau + \delta) \frac{d}{d\delta} \Sigma(\tau, \tau + \delta)$ and $2\Sigma(\tau, \tau + \delta)$ would equal zero when evaluated at $\delta = 0$ (see equation 72), which would leave the limit undefined.

Re-arranging equation 92 gives:

$$\left[\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} \right]^2 = \frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \tau + \delta)]^2 \right\} \quad (93)$$

and therefore:

$$\omega(\tau) = \sqrt{\frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \tau + \delta)]^2 \right\}} \quad (94)$$

Referring to equation 31, $\lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \tau + \delta)]^2 \right\}$ in equation 94 may be evaluated by calculating the limits of the double derivatives of the functions $[G(\kappa_n, \delta)]^2$ and $G(\kappa_m, \delta) G(\kappa_n, \delta)$ and then substituting those results into equation 31. That is:

$$\begin{aligned} \frac{d^2}{d\delta^2} [G(\kappa_n, \delta)]^2 &= \frac{d}{d\delta} \left[\frac{d}{d\delta} [G(\kappa_n, \delta)]^2 \right] \\ \langle \text{Chain rule} \rangle &= \frac{d}{d\delta} [2G(\kappa_n, \delta)] \exp(-\kappa_n \delta) \\ &= 2 \frac{d}{d\delta} G(\kappa_n, \delta) \cdot \exp(-\kappa_n \delta) + 2G(\kappa_n, \delta) \frac{d}{d\delta} \exp(-\kappa_n \delta) \\ &= 2 \exp(-\kappa_n \delta) \cdot \exp(-\kappa_n \delta) + 2G(\kappa_n, \delta) \cdot -\kappa_n \exp(-\kappa_n \delta) \\ &= 2 \exp(-2\kappa_n \delta) - 2G(\kappa_n, \delta) \cdot \kappa_n \exp(-\kappa_n \delta) \end{aligned} \quad (95)$$

which has a limit:

$$\lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [G(\kappa_n, \delta)]^2 \right\} = 2 - 0 = 2 \quad (96)$$

and:

$$\begin{aligned} \frac{d^2}{d\delta^2} [G(\kappa_m, \delta) G(\kappa_n, \delta)] &= \frac{d}{d\delta} \left[\frac{d}{d\delta} [G(\kappa_m, \delta) G(\kappa_n, \delta)] \right] \\ &= \frac{d}{d\delta} [\exp(-\kappa_m \delta) G(\kappa_n, \delta) + G(\kappa_m, \delta) \exp(-\kappa_n \delta)] \\ &= \frac{d}{d\delta} \exp(-\kappa_m \delta) \cdot G(\kappa_n, \delta) + \exp(-\kappa_m \delta) \frac{d}{d\delta} G(\kappa_n, \delta) \\ &\quad + \frac{d}{d\delta} G(\kappa_m, \delta) \cdot \exp(-\kappa_n \delta) \\ &\quad + G(\kappa_m, \delta) \frac{d}{d\delta} \exp(-\kappa_n \delta) \\ &= -\kappa_m \exp(-\kappa_m \delta) G(\kappa_n, \delta) + \exp(-\kappa_m \delta) \exp(-\kappa_n \delta) \\ &\quad + \exp(-\kappa_m \delta) \exp(-\kappa_n \delta) \\ &\quad + G(\kappa_m, \delta) \cdot -\kappa_n \exp(-\kappa_n \delta) \\ &= -\kappa_m \exp(-\kappa_m \delta) G(\kappa_n, \delta) + 2 \exp(-\{\kappa_m + \kappa_n\} \delta) \\ &\quad - G(\kappa_m, \delta) \kappa_n \exp(-\kappa_n \delta) \end{aligned} \quad (97)$$

which has a limit:

$$\lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [G(\kappa_m, \delta) G(\kappa_n, \delta)] \right\} = 0 + 2 + 0 = 2 \quad (98)$$

Substituting the two limit results into equation 31 $\lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \tau + \delta)]^2 \right\}$ gives:

$$\begin{aligned} \lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\cdot)]^2 \right\} &= \sum_{n=1}^N 2\sigma_n^2 [G(\kappa_n, 2\tau)]^2 \\ &\quad + 2 \sum_{m=1}^N \sum_{n=m+1}^N 2\rho_{mn} \sigma_m \sigma_n G(\kappa_m + \kappa_n, 2\tau) \end{aligned} \quad (99)$$

and substituting the preceding result into equation 94 gives:

$$\omega(\tau) = \sqrt{\sum_{n=1}^N \sigma_n^2 [G(2\kappa_n, \tau)]^2 + 2 \sum_{m=1}^N \sum_{n=m+1}^N \rho_{mn} \sigma_m \sigma_n G(\kappa_m + \kappa_n, \tau)} \quad (100)$$

C Generic GATSM to HJM model

This appendix details how the HJM model in sections 4.2 may be replicated within the generic GATSM framework.

I first set $N = 2$ in equation 36, and then set $\mu_1 = \mu_2 = 0$. The resulting interim expression is:

$$f(t, \tau) = -\frac{1}{2} \text{Tr} [\Theta(\tau) \Psi] + \sum_{n=1}^2 s_n(t) \cdot \exp(-\kappa_n \tau) \quad (101)$$

Uncorrelated factor innovations are replicated with $\rho_{12} = 0$, so the matrix $\Theta(\tau)$ has only diagonal elements $\Theta_{ii}(\tau) = \sigma_i^2 \kappa_i^2 [G(\kappa_i, \tau)]^2$. The trace evaluation $\text{Tr}[\cdot]$ in equation 101 is therefore:

$$\begin{aligned} \text{Tr}[\cdot] &= \text{Tr} \left\{ \begin{bmatrix} \sigma_1^2 \cdot \kappa_1^2 [G(\kappa_1, \tau)]^2 & 0 \\ 0 & \sigma_2^2 \cdot \kappa_2^2 [G(\kappa_2, \tau)]^2 \end{bmatrix} \begin{bmatrix} \frac{1}{\kappa_1^2} & \frac{1}{\kappa_1 \kappa_2} \\ \frac{1}{\kappa_1 \kappa_2} & \frac{1}{\kappa_2^2} \end{bmatrix} \right\} \\ &= \text{Tr} \left\{ \begin{bmatrix} \sigma_1^2 \cdot [G(\kappa_1, \tau)]^2 & \text{n/a} \\ \text{n/a} & \sigma_2^2 \cdot [G(\kappa_2, \tau)]^2 \end{bmatrix} \right\} \\ &= \sigma_1^2 \cdot [G(\kappa_1, \tau)]^2 + \sigma_2^2 \cdot [G(\kappa_2, \tau)]^2 \end{aligned} \quad (102)$$

giving the following expression for the forward rate:

$$\begin{aligned} f(t, \tau) &= s_1(t) \cdot \exp(-\kappa_1 \tau) + s_2(t) \cdot \exp(-\kappa_2 \tau) \\ &\quad - \sigma_1^2 \cdot \frac{1}{2} [G(\kappa_1, \tau)]^2 - \sigma_2^2 \cdot \frac{1}{2} [G(\kappa_2, \tau)]^2 \end{aligned} \quad (103)$$

The expression for annualized instantaneous volatility from equation 39 will have no cross-correlation terms, given $\rho_{12} = 0$. Therefore:

$$\omega(\tau) = \sqrt{\sigma_1^2 \cdot [G(2\kappa_1, \tau)]^2 + \sigma_2^2 \cdot [G(2\kappa_2, \tau)]^2} \quad (104)$$

Both $f(t, \tau)$ and $\omega(\tau)$ can be made more parsimonious by taking the limit as $\kappa_1 \rightarrow 0$, which reproduces the HJM model. Specifically: $\lim_{\kappa_1 \rightarrow 0} \exp(-\kappa_1 \tau) = 1$,

$\lim_{\kappa_1 \rightarrow 0} \left\{ \frac{1}{2} [G(\kappa_1, \tau)]^2 \right\} = \frac{1}{2} \tau^2$, and $\lim_{\kappa_1 \rightarrow 0} [G(2\kappa_1, \tau)]^2 = \tau^2$, where the latter two results follow from the following limit evaluation:

$$\begin{aligned}
\lim_{\kappa_1 \rightarrow 0} \{ [G(x, \tau)]^2 \} &= \left[\lim_{\kappa_1 \rightarrow 0} G(x, \tau) \right]^2 \\
&= \left[\lim_{x \rightarrow 0} \left\{ \frac{1 - \exp(-x\tau)}{x} \right\} \right]^2 \\
\langle \text{L'Hopital's rule} \rangle &= \left[\frac{\lim_{x \rightarrow 0} \left\{ \frac{d}{dx} [1 - \exp(-x\tau)] \right\}}{\lim_{x \rightarrow 0} \left\{ \frac{d}{dx} x \right\}} \right]^2 \\
&= \left[\lim_{x \rightarrow 0} \{ \tau \exp(-x\tau) \} \right]^2 \\
&= \tau^2
\end{aligned} \tag{105}$$

Finally, I set $\kappa_2 = \lambda$ for notational convenience. The HJM model is therefore:

$$f(t, \tau) = s_1(t) + s_2(t) \cdot \exp(-\lambda\tau) - \sigma_1^2 \cdot \frac{1}{2} \tau^2 - \sigma_2^2 \cdot \frac{1}{2} [G(\lambda, \tau)]^2$$

with:

$$\omega(\tau) = \sqrt{\sigma_1^2 \cdot \tau^2 + \sigma_2^2 \cdot [G(2\lambda, \tau)]^2} \tag{106}$$