RESERVE BANK

OF NEW ZEALAND

DP2012/02

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

Leo Krippner

March 2012

JEL classification: E43, G12, G13

www.rbnz.govt.nz/research/discusspapers/

Discussion Paper Series

ISSN 1177-7567

DP2012/02

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

Leo Krippner†

Abstract

With nominal interest rates near the zero lower bound (ZLB) in many major economies, it is theoretically untenable to apply Gaussian affine term structure models (GATSMs) while ignoring their inherent material probabilities of negative interest rates. I propose correcting that deficiency by adjusting the entire GATSM term structure with an explicit function of maturity that represents the optionality associated with the present and future availability of physical currency. 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 estimation with U.S. term structure data indicates that the ZLB-GATSM "shadow short rate" provides a useful gauge of the stance of monetary policy; in particular becoming negative when the U.S. policy rate reached the ZLB in late 2008, and moving more negative with subsequent unconventional monetary policy easings.

* The Reserve Bank of New Zealand's discussion paper series is externally refereed. The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Reserve Bank of New Zealand. The author thanks Katy Bergstrom, Iris Claus, Toby Daglish, Francis Diebold, Pedro Gomis, Michelle Lewis, Anella Munro, Glenn Rudebusch, Peter Phillips, Christie Smith, Daniel Thornton, Christopher Waller, participants at the 2011 Reserve Bank of New Zealand conference, a 2011 Bundesbank seminar, and the 2012 New Zealand Econometrics Study Group meeting for helpful comments.

† Economics Department, Reserve Bank of New Zealand, 2 The Terrace, PO Box 2498, Wellington, New Zealand. E-mail address: leo.krippner@rbnz.govt.nz.

ISSN 1177-7567

1 Introduction

In this article I propose a 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 address the inherent theoretical deficiency of negative interest rates in GATSMs, and a practical, topical motivation is to provide an indicator of the stance of monetary policy in the presence of the ZLB. I expand on both motivations subsequently below. At the same time, I seek to preserve in the framework 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 both closed-form analytic solutions for pricing standard interest rate instruments (e.g. bonds and options) and multivariate-normal transition densities for the state variables. Those features make GATSMs straightforward to estimate and apply relative to other term structure models. As such, Hamilton and Wu (2010) introduces GATSMs as "the basic workhorse in macroeconomics and finance" and notes seven recent examples of their application. Rudebusch (2010) surveys GATSM applications in macrofinance.[1]

However, it is well acknowledged that GATSMs cannot be theoretically consistent in the "real world" where physical currency is available (e.g. see Piazzesi (2010) p. 716). The inconsistency arises because any unconstrained Gaussian process for short rate dynamics technically implies non-zero probabilities of negative interest rates for all maturities on the term structure. 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 that known inconsistency, GATSMs are often applied with the assumption (as discussed in Piazzesi (2010) p. 716, but typically 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 may have been the case over history, when interest rate levels relative to their typical standard deviation of unanticipated changes remained well above zero, such an assumption is obviously untenable in several major developed economies at the time of writing. For example, 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 (and each of those countries have also engaged in unconventional monetary policy easings due to the ZLB constraint). Moreover, the levels of short- and medium-maturity interest rates in those countries also lie well within their typical standard deviations of unanticipated changes.

The material probability of negative interest rates in GATSMs in near-zero interest

[1] Hamilton and Wu (2010) proposes a reliable method for estimating GATSMs, and Joslin, Singleton, and Zhu (2011) provides another approach. Arbitrage-free versions of the Nelson and Siegel (1987) class are a sub-class of GATSMs that are also empirically reliable. (Their nesting within the GATSM class is clearly illustrated in Christensen, Diebold, and Rudebusch (2009, 2011), where arbitrage-free Nelson and Siegel (1987) models are derived via GATSM specifications that reproduce Nelson and Siegel (1987) factor loadings.)

rate environments in turn implies model mis-specification relative to the data being modeled. 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 mis-specification 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.

A straightforward modification to GATSMs that eliminates negative interest rates is the Black (1995) framework,[2] which is based on the observation that physical currency effectively provides an option against negative interest rates at each point in time. Specifically, the ZLB short rate may be defined as $\underline{\mathrm{r}}(t) = \max\{\mathrm{r}(t), 0\}$, where $\mathrm{r}(t)$ is the "shadow short rate" that is free to evolve with negative and positive values.

Unfortunately, Black-GATSM implementations (i.e. using GATSMs to define the shadow short rate $\mathrm{r}(t)$ within the Black (1995) framework) result in models with limited tractability. For example, even the simplest Black-GATSM, the one-factor Black-Vasicek (1977) model by Gorovoi and Linetsky (2004), does not result in a closed-form analytic solution. Specifically, it requires the numerical evaluation of relatively complex functions (e.g. Weber-Hermite parabolic cylinder functions). In addition, the Gorovoi and Linetsky (2004) approach does not appear to generalize to multiple factors.[3] Bomfim (2003), Ueno, Baba, and Sakurai (2006), and Ichiue and Ueno (2007) respectively use the purely numerical methods of finite-difference grids, Monte Carlo simulations, and interest rate lattices to illustrate and/or estimate two-factor Black-GATSMs.[4] However, the complexity of those numerical methods increases to the order of the power of the number of factors. Therefore, while Black-GATSMs have arbitrary flexibility in principle, in practice their rapid decline in tractability would preclude the full flexibility offered by larger multifactor GATSMs.

The ZLB framework I propose in this article is conceptually similar to Black (1995), in that it is based on the optionality provided by physical currency, but it is distinctly different in the following respect: I ensure the entire ZLB forward rate curve remains

---

[2] Of course, a prevalent literature has evolved over several decades using non-Gaussian dynamics designed to avoid negative interest rates in term structure models; e.g. James and Webber (2000) pp. 226-233 discusses a variety of positive interest rate models. Given my focus is on Gaussian models, I do not discuss positive interest rate models further beyond the following comments in this context of this article: (1) such models lose the potential information provided by the "shadow short rate" and "shadow term structure", as will be discussed subsequently in the main text; and (2) such models inevitably result in limited flexibility and/or tractability relative to GATSMs. For example on the latter point, 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 (e.g. see Piazzesi (2010) p. 727).

[3] See 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 term structure.

[4] The Bomfim (2003) application is to United States data, as I subsequently discuss in section 4.3. The other applications are to the Japanese term structure. An analogous Japanese application is the two-factor Black-quadratic-Gaussian-model from Kim and Singleton (2011), which nests the two-factor Black-GATSM as a special case.

non-negative at each point in time by adding to the shadow forward rate curve an explicit function of maturity that represents the option effect from the present and future availability of physical currency. To highlight just the essence of the distinction at this stage,[5] the Black (1995) framework effectively allows for an option effect of $\max\{-\mathrm{r}(t), 0\}$ (which is positive if $\mathrm{r}(t)$ is negative, and zero otherwise) at each point in time. While that mechanism also imposes the ZLB constraint, i.e. $\mathrm{r}(t) + \max\{-\mathrm{r}(t), 0\} = \max\{0, \mathrm{r}(t)\} = \underline{\mathrm{r}}(t)$, the function of maturity representing the future stream of physical currency optionality is left implicit.

My proposed ZLB framework in the Gaussian context uses the closed-form expression for GATSM call options on bonds to derive the option effect as a function of time to maturity. Adding that to the shadow-GATSM forward rate curve results in a ZLB-GATSM forward rate curve with a simple closed-form analytic expression, involving just exponential functions and the univariate cumulative normal distribution. ZLB-GATSM interest rates and ZLB-GATSM bond prices are obtained using standard term structure relationships; respectively, the mean of the integral of ZLB-GATSM forward rates over time to maturity and the exponential of that (negated) integral. The integral is necessarily numerical in the Gaussian context, but its evaluation is elementary due to the nature of the ZLB-GATSM forward rate expression.

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 a selection of shadow-GATSM forward and interest rate curves. The second exercise is an application to U.S. term structure data and illustrates that model-implied shadow short rates move 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. Those movements indicate that the ZLB-GATSM framework provides a means for routinely distilling a quantitative indicator of the stance of monetary policy from term structure data when interest rates are materially constrained by the ZLB.

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 based on bond and bond option prices, and then discusses the differences between my framework and the framework of Black (1995). Section 3 derives the generic ZLB-GATSM term structure. Section 4 contains the illustrative exercises already noted earlier, including a brief discussion of recent related literature. Section 5 briefly discusses potential extensions of the ZLB-GATSM framework introduced in this article, including aspects relating to a full empirical estimation and a thorough exploration of the resulting ZLB-GATSM shadow short rate and the shadow term structure as indicators of the stance of monetary policy. Section 6 concludes.

2 A general ZLB framework

In this section, I outline my approach to constructing ZLB term structures without referring to any particular model of the term structure. Section 2.1 introduces the concept of the availability of physical currency effectively providing investors with a

[5] Section 2.4 precisely defines the distinction, in light of appropriate notation, between my proposed ZLB framework and the Black (1995) framework.

(European) call option on a bond, in addition to the bond investment itself.[6] Section 2.2 extends that optionality concept from the current point in time to future time horizons. Section 2.3 uses the results from section 2.2 to derive a general ZLB term structure in terms of bond and option prices. I discuss in section 2.4 how my proposed ZLB framework differs from the Black (1995) framework.

2.1 Physical currency over a single time-step

Physical currency may be viewed as a bond with a price of 1 and an associated interest rate of zero. To illustrate this, I introduce a finite-step shadow bond that settles for price $\mathrm{P}(t,\delta)$ at time t and pays 1 at time $t+\delta$, where $\delta > 0$ is a (small) finite-step representing the time to maturity.[7] The annualized rate of return over the finite-step on a continuously-compounding basis is the finite-step shadow short rate $\mathrm{R}(t,\delta)$, i.e.:

$$\mathrm{R}(t,\delta) = \frac{1}{\delta}\log\left[\frac{1}{\mathrm{P}(t,\delta)}\right] \tag{1}$$

Clearly, if $\mathrm{P}(t,\delta) = 1$, then $\mathrm{R}(t,\delta) = 0$, and so the return equals that offered by physical currency.

The availability of physical currency is therefore equivalent to finite-step shadow bonds with a price $\mathrm{P}(t,\delta) = 1$ 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 $\mathrm{P}(t,\delta)$ is less than 1, then $\mathrm{R}(t,\delta) > 0$ and so investors will choose to hold $\mathrm{P}(t,\delta)$ rather than substituting physical currency at $\mathrm{P}(t,\delta) = 1$. Alternatively, if the prevailing market price of $\mathrm{P}(t,\delta)$ is greater than 1, then $\mathrm{R}(t,\delta) < 0$ and so investors will choose to maximize their return by holding physical currency at $\mathrm{P}(t,\delta) = 1$.

In summary, investors will choose $\underline{\mathrm{P}}(t,\delta) = \min\{1,\mathrm{P}(t,\delta)\}$, where $\underline{\mathrm{P}}(t,\delta)$ is my notation for the finite-step ZLB bond. $\underline{\mathrm{P}}(t,\delta)$ may be expressed more conveniently as the sum of the finite-step shadow bond $\mathrm{P}(t,\delta)$ and the payoff for a call option on $\mathrm{P}(t,\delta)$, i.e.:

$$\begin{aligned} \underline{\mathrm{P}}(t,\delta) &= \min\{1,\mathrm{P}(t,\delta)\} \\ &= \mathrm{P}(t,\delta) + \min\{1-\mathrm{P}(t,\delta),0\} \\ &= \mathrm{P}(t,\delta) - \max\{\mathrm{P}(t,\delta)-1,0\} \\ &= \mathrm{P}(t,\delta) - \mathrm{C}(t,0,\delta) \end{aligned} \tag{2}$$

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

The optionality arising from the availability of physical currency over the finite step δ establishes the lower bound of zero for the finite-step ZLB short rate $\underline{\mathrm{R}}(t,\delta)$.

[6] I typically take the perspective of investors, but all comments and calculations from the investors' perspective can be reversed to give the borrowers' perspective.

[7] For example, overnight or on-call deposits (often colloquially referred to as "cash" or "liquidity") are effectively finite-step bonds with a single day to maturity.

Explicitly:

$$\begin{aligned}
\underline{\mathrm{R}}(t,\delta) &= \frac{1}{\delta}\log\left[\frac{1}{\underline{\mathrm{P}}(t,\delta)}\right] \\
&= -\frac{1}{\delta}\log\left[\mathrm{P}(t,\delta)-\mathrm{C}(t,0,\delta)\right] \\
&= -\frac{1}{\delta}\log\left[\min\left\{1,\mathrm{P}(t,\delta)\right\}\right] \\
&= \left[\max\left\{-\frac{1}{\delta}\log[1],-\frac{1}{\delta}\log\left[\mathrm{P}(t,\delta)\right]\right\}\right] \\
&= \max\left\{0,\mathrm{R}(t,\delta)\right\} \qquad (3)
\end{aligned}$$

Note that taking the infinitesimal limit of δ would reproduce the Black (1995) concept of physical currency providing an option on the shadow short rate; heuristically, $\lim_{\delta\to 0}\underline{\mathrm{R}}(t,\delta) = \max\left[\lim_{\delta\to 0}\mathrm{R}(t,\delta),0\right]$, so $\underline{\mathrm{r}}(t) = \max\left[\mathrm{r}(t),0\right]$. Section 2.3 contains a more formal exposition of that infinitesimal limit in the context of defining the continuous-time term structure for the general ZLB framework.

2.2 Physical currency over multiple time-steps

Beyond the first finite-step, investors also know that the availability of physical currency at any point of time in the future will always offer them the choice between holding physical currency or investing in shadow bonds at their prevailing market price. In notation, I introduce a positive quantity τ to represent a future horizon from time t, so future time may be represented as $t+\tau$.[8] The multiple-finite-step ZLB bond is therefore defined as:

$$\underline{\mathrm{P}}(t+\tau,\delta) = \min\left\{1,\mathrm{P}(t+\tau,\delta)\right\} \qquad (4)$$

While the optionality inherent in $\underline{\mathrm{P}}(t+\tau,\delta)$ is analogous to that for the first-finite-step ZLB bond $\underline{\mathrm{P}}(t,\delta)$, the difference is that $\underline{\mathrm{P}}(t,\delta)$ is a non-contingent quantity (because the choice between holding the bond or physical currency is already known at time t), while $\underline{\mathrm{P}}(t+\tau,\delta)$ is a contingent quantity (because the choice between holding the bond or physical currency is unknown at time t, being dependent on the single-finite-step shadow bond price $\mathrm{P}(t+\tau,\delta)$ at the future time $t+\tau$). Nevertheless, $\underline{\mathrm{P}}(t+\tau,\delta)$ provides a terminal boundary condition that may be used in conjunction with an appropriate pricing mechanism to obtain the multiple-finite-step ZLB bond price $\underline{\mathrm{P}}(t,\tau+\delta)$. An explicit mechanism is not required for the present discussion, but section 2.4.2 introduces a generic expression for security pricing in the context of comparing my proposed ZLB framework to the Black (1995) framework.

Analogous to the first-finite-step ZLB bond $\underline{\mathrm{P}}(t,\delta)$, the multiple-finite-step ZLB bond $\underline{\mathrm{P}}(t,\tau+\delta)$ may also be more conveniently expressed as the sum of the multiple-finite-step shadow bond $\mathrm{P}(t,\tau+\delta)$ and a call option on $\mathrm{P}(t,\tau+\delta)$. To establish that sum, I proceed by first separating the boundary condition for the ZLB bond into the

[8] For the purpose of transparent exposition within a discrete time formulation, I assume at this stage that t and τ are integer multiples of δ. However, the finite-step δ is the key component for the discrete time formulation, while t and τ could be regarded as continuous variables.

boundary condition for a shadow bond and a put option, i.e.:

$$\begin{aligned}\underline{\mathrm{P}}(t+\tau,\delta) &= \min\{1,\mathrm{P}(t+\tau,\delta)\} \\ &= 1+\min\{0,\mathrm{P}(t+\tau,\delta)-1\} \\ &= 1-\max\{0,1-\mathrm{P}(t+\tau,\delta)\}\end{aligned} \tag{5}$$

where 1 is the boundary condition for the multiple-finite-step shadow bond $\mathrm{P}(t,\tau)$ maturing at time $t+\tau$ (i.e. $\mathrm{P}(t+\tau,0)=1$), and $\max\{0,1-\mathrm{P}(t+\tau,\delta)\}$ is the boundary condition for a put option $\mathrm{Q}(t,\tau,\tau+\delta)$ expiring at time $t+\tau$ with a strike price of 1.[9] Note that the underlying security for $\mathrm{Q}(t,\tau,\tau+\delta)$ at expiry is the single-finite-step shadow bond $\mathrm{P}(t+\tau,\delta)$, i.e. $\mathrm{Q}(t+\tau,0,\delta)=\max\{0,1-\mathrm{P}(t+\tau,\delta)\}$, while the underlying security at time t is the forward price of that bond.

The boundary condition $\mathrm{P}(t+\tau,0)=1$ obtains the shadow bond price $\mathrm{P}(t,\tau)$, while the boundary condition $-\max\{0,1-\mathrm{P}(t+\tau,\delta)\}$ obtains the negated put option price $-\mathrm{Q}(t,\tau,\tau+\delta)$. Therefore, the solution for the ZLB bond price is:

$$\underline{\mathrm{P}}(t,\tau+\delta)=\mathrm{P}(t,\tau)-\mathrm{Q}(t,\tau,\tau+\delta) \tag{6}$$

The right-hand side of equation 6 may be re-expressed in terms of $\mathrm{P}(t,\tau+\delta)$ and $\mathrm{C}(t,\tau,\tau+\delta)$ using the put-call parity relationship for forward bonds and bond option prices, as noted in Chen (1995) p. 363, with a strike price of 1. That is, $\mathrm{C}(t,\tau,\tau+\delta)-\mathrm{Q}(t,\tau,\tau+\delta)=\mathrm{P}(t,\tau+\delta)-\mathrm{P}(t,\tau)$,[10] and so:

$$\mathrm{P}(t,\tau)-\mathrm{Q}(t,\tau,\tau+\delta)=\mathrm{P}(t,\tau+\delta)-\mathrm{C}(t,\tau,\tau+\delta) \tag{7}$$

The ZLB bond price therefore becomes:

$$\underline{\mathrm{P}}(t,\tau+\delta)=\mathrm{P}(t,\tau+\delta)-\mathrm{C}(t,\tau,\tau+\delta) \tag{8}$$

Finally, analogous to the first-finite-step ZLB short rate $\underline{\mathrm{R}}(t,\delta)$, the optionality arising from the future availability of physical currency establishes that each realized future single-finite-step ZLB short rate $\underline{\mathrm{R}}(t+\tau,\delta)$ will respect the ZLB. Explicitly:

$$\begin{aligned}\underline{\mathrm{R}}(t+\tau,\delta) &= \frac{1}{\delta}\log\left[\frac{1}{\underline{\mathrm{P}}(t+\tau,\delta)}\right] \\ &= -\frac{1}{\delta}\log\left[\mathrm{P}(t+\tau,\delta)-\mathrm{C}(t+\tau,0,\delta)\right] \\ &= -\frac{1}{\delta}\log\left[\min\{1,\mathrm{P}(t+\tau,\delta)\}\right] \\ &= \left[\max\left\{-\frac{1}{\delta}\log[1],-\frac{1}{\delta}\log[\mathrm{P}(t+\tau,\delta)]\right\}\right] \\ &= \max\{0,\mathrm{R}(t+\tau,\delta)\}\end{aligned} \tag{9}$$

[9] Expressions for put option prices are often denoted $\mathrm{P}(\cdot)$, but I have used the notation $\mathrm{Q}(\cdot)$ to avoid any confusion with my notation $\mathrm{P}(\cdot)$ for bond prices.

[10] The standard put-call parity relationship equates the forward price to the call option price less the put option price, or $F=C-Q$ (see, for example, Haug (2007) p. 18 or the original reference therein to Nelson (1904)). For forward bonds and options with a strike price of 1, $F=1=\mathrm{P}(t,\tau+\delta)/\mathrm{P}(t,\tau)$, so $F=\mathrm{P}(t,\tau+\delta)-\mathrm{P}(t,\tau)$.

2.3 A general ZLB term structure

While the finite-step securities in sections 2.1 and 2.2 could be used directly to establish a discrete-time term structure model, it is almost invariably more convenient to work with term structure models in continuous time. For example, in the ZLB-GATSM framework that follows in section 3, the continuous-time expressions are much more parsimonious and computationally convenient than their discrete-time counterparts would be.

To obtain the continuous-time term structure for the general ZLB framework, I begin with the standard term structure relationship relating (continuously compounding) ZLB forward rates to ZLB bond prices, i.e.:[11]

$$\underline{\mathrm{f}}(t,\tau) = -\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\underline{\mathrm{P}}(t,\tau)\right] \tag{10}$$

and I obtain the required derivative via $\mathrm{P}(t,\tau+\delta)$ in the infinitesimal limit of δ, i.e.:

$$-\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\underline{\mathrm{P}}(t,\tau)\right] = \lim_{\delta\to 0}\left\{-\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\log\left[\underline{\mathrm{P}}(t,\tau+\delta)\right]\right\} \tag{11}$$

Appendix A.1 contains the details of this derivation. The resulting expression for the ZLB forward rate is:

$$\underline{\mathrm{f}}(t,\tau) = \mathrm{f}(t,\tau) + \mathrm{z}(t,\tau) \tag{12}$$

where $\mathrm{f}(t,\tau)$ is the shadow forward rate curve and $\mathrm{z}(t,\tau)$ is the option effect:

$$\mathrm{z}(t,\tau) = \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{C}(t,\tau,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\} \tag{13}$$

The ZLB forward rate $\underline{\mathrm{f}}(t,\tau)$ establishes the ZLB interest rate $\underline{\mathrm{r}}(t,\tau)$ and the ZLB bond price $\underline{\mathrm{P}}(t,\tau)$ via standard term structure relationships, i.e.:

$$\underline{\mathrm{r}}(t,\tau) = \frac{1}{\tau}\int_0^{\tau}\underline{\mathrm{f}}(t,\upsilon)\,\mathrm{d}\upsilon \tag{14}$$

where υ is a dummy integration variable for horizon/time to maturity, and:

$$\begin{aligned}\underline{\mathrm{P}}(t,\tau) &= \exp\left[-\int_0^{\tau}\underline{\mathrm{f}}(t,\upsilon)\,\mathrm{d}\upsilon\right]\\ &= \exp\left[-\tau\cdot\underline{\mathrm{r}}(t,\tau)\right]\end{aligned} \tag{15}$$

It is worthwhile explicitly establishing that the ZLB short rate $\underline{\mathrm{r}}(t)$ from the general ZLB term structure will always be a well-defined and non-negative quantity. Appendix A.2 contains the details, which I summarize as follows:

$$\begin{aligned}\underline{\mathrm{r}}(t) &= \lim_{\tau\to 0}\underline{\mathrm{f}}(t,\tau)\\ &= \mathrm{f}(t,0) + \mathrm{z}(t,0)\\ &= \mathrm{r}(t) + \max\left\{-\mathrm{r}(t),0\right\}\\ &= \max\left\{\mathrm{r}(t),0\right\}\end{aligned} \tag{16}$$

[11] See, for example, Filipović (2009) p. 7 or James and Webber (2000) chapter 3 for this relationship and other standard term structure relationships that I subsequently refer to and use in this article.

2.4 Comparing the Black (1995) and ZLB frameworks

Section 2.4.1 establishes by example that the Black (1995) framework and my proposed ZLB framework are not equivalent. In section 2.4.2, I discuss the difference between the two frameworks from a theoretical perspective.

2.4.1 Numerical comparison

It is straightforward to establish by example that the Black (1995) and ZLB frameworks are not equivalent. Specifically, using an identical model for the shadow short rate, I compare the numerical results from my ZLB framework with those of the Black (1995) framework.

For the comparison, tables 1 and 2 respectively present what I call the BGL-GATSM(1) results from table 6.1 in Gorovoi and Linetsky (2004), which are based on a one-factor GATSM/Vasicek (1977) model within the Black (1995) framework.[12] The precise GATSM(1) specification for the shadow short rate is the shadow-GATSM(2) model subsequently specified in section 4.1 of the present article with the state variable $\mathrm{r}(t) = s_2(t)$ and the parameters $\lambda = 0.1$, $\mu_2 = 0.01$, $\sigma_2 = 0.02$, and $\gamma_2 = 0$.[13] Gorovoi and Linetsky (2004) uses initial values of $\mathrm{r}(t) = 0$ and $\mathrm{r}(t) = 0.01$ (i.e. 1 percent) for the shadow short rate.

I then use the same GATSM(1) specification within my ZLB-GATSM framework to evaluate the corresponding ZLB-GATSM(1) interest rates (the approach is subsequently outlined for the ZLB-GATSM(2) in section 4.1). The results are shown in tables 1 and 2, along with the differences between the Black (1995) and ZLB framework results.

Table 1:
Interest rates (in percentage points) with $\mathrm{r}(t) = 0$ percent

Model\time to maturity (years)	1	5	10	30
BGL-GATSM(1)	0.539	1.106	1.378	1.634
ZLB-GATSM(1)	0.538	1.084	1.314	1.422
BGL less ZLB	0.001	0.022	0.064	0.211
GATSM(1)	0.042	0.097	0.032	-0.382

[12] Note that, to facilitate comparison across the different times to maturity, I have converted the Gorovoi and Linetsky (2004) bond price results to their interest rate equivalents using the standard term structure relationship $\mathrm{r}(t,\tau) = -\frac{1}{\tau}\log[\mathrm{P}(t,\tau)]$. The bond prices from Gorovoi and Linetsky (2004) are accurate to five decimal places. That accuracy translates to the interest rates in tables 1 and 2 being accurate to at least three decimal places (i.e. 0.001 percentage points or 0.1 basis points) in all cases. I have also reproduced the Gorovoi and Linetsky (2004) results independently via Monte-Carlo simulation.

[13] The state variable $s_1(t)$ and its associated parameters in the GATSM(2) are, of course, all set to zero.

Table 2: Interest rates (percentage points) with $\mathrm{r}(t) = 1$ percent				
Model\time to maturity (years)	1	5	10	30
BGL-GATSM(1)	1.178	1.570	1.731	1.795
ZLB-GATSM(1)	1.177	1.552	1.673	1.592
BGL less ZLB	0.001	0.019	0.058	0.204
GATSM(1)	0.994	0.884	0.664	-0.066

While the respective results are very similar for the 1-year maturity, the differences are sufficiently large for the other maturities considered (e.g. a maximum of 0.211 percentage points at the maturity of 30 years in table 1, relative to the significance of 0.001 percentage points) to confirm that $\underline{\mathrm{r}}(t,\tau) \neq \underline{\mathrm{r}}_{\mathrm{Black}}(t,\tau)$, or equivalently $\underline{\mathrm{P}}(t,\tau) \neq \underline{\mathrm{P}}_{\mathrm{Black}}(t,\tau)$.

As an aside to the main point of this section, but relevant for practically illustrating the theoretical deficiency of GATSMs in low interest rate environments, note that the interest rates for the GATSM(1) are always substantially below their counterparts that use a ZLB mechanism. Indeed, the relatively low mean reversion of $\lambda = 0.1$ in these particular examples results in negative interest rates for the 30-year maturity (i.e. bond prices greater than 1), which is a phenomenon I return to discuss in the context of the GATSM(2) examples of section 4.2.

2.4.2 Theoretical comparison

The natural follow-up question to section 2.4.1 is: how/where does the difference between the two frameworks arise?

The first point to note in response is that the two frameworks have identical short rates under the physical or $\mathbb{P}$ measure. That is straightforward to establish by simply noting that implementations of the Black (1995) framework are based on short rate diffusion processes defined as $\underline{\mathrm{r}}_{\mathrm{Black}}(t) = \max\{\mathrm{r}(t), 0\}$, while equation 16 establishes that the short rate process associated with my proposed ZLB framework is $\underline{\mathrm{r}}(t) = \max\{\mathrm{r}(t), 0\}$. Therefore, realized short rates for a given shadow short rate process are identical in the two frameworks; i.e. $\underline{\mathrm{r}}(t) = \underline{\mathrm{r}}_{\mathrm{Black}}(t)$.

The result $\underline{\mathrm{P}}(t,\tau) \neq \underline{\mathrm{P}}_{\mathrm{Black}}(t,\tau)$ from section 2.4.1 along with $\underline{\mathrm{r}}(t) = \underline{\mathrm{r}}_{\mathrm{Black}}(t)$ under the $\mathbb{P}$ measure implies that the Black (1995) and ZLB frameworks must have different short rate diffusion processes under the risk-neutral $\mathbb{Q}$ measure. To make that proposition more transparent, I introduce the standard generic expression that defines security prices via their expected discounted cashflows; i.e.:[14]

$$\mathrm{X}(t,\tau) = \mathbb{E}_t^{\mathbb{Q}}\left\{\exp\left(-\int_0^\tau \check{\mathrm{r}}(t+v)\,\mathrm{d}v\right)\cdot \mathrm{A}(t+\tau)\right\} \tag{17}$$

where $\mathrm{X}(t,\tau)$ is the security to be priced (based on a given set of state variables $s(t)$ and parameters $\mathbb{B}$), $\mathbb{E}_t^{\mathbb{Q}}$ is the expectations operator under the $\mathbb{Q}$ measure, $\mathrm{A}(t+\tau)$

[14] See, for example, the generic option price expression from Filipović (2009) p. 109, which is given in appendix A.1 of the present article. Filipović (2009) chapter 4 and James and Webber (2000) chapter 4 are two examples of many texts that discuss the standard framework of risk-neutral security pricing.

is the terminal cashflow for the security $\mathrm{X}(t,\tau)$, $\check{\mathrm{r}}(t+\upsilon)$ is the $\mathbb{Q}$-measure short rate diffusion process, which is used to obtain the discount factor $\exp\left(-\int_0^\tau \check{\mathrm{r}}\left(\upsilon\right)\mathrm{d}\upsilon\right)$, and υ is a dummy integration variable for future time relative to the current time t.

Bond prices in the Black (1995) framework, which I denote $\underline{\mathrm{P}}_{\mathrm{Black}}\left(t,\tau\right)$, are defined directly via equation 17 using $\check{\mathrm{r}}(t+\upsilon) = \max\left\{\mathrm{r}\left(t+\upsilon\right), 0\right\} = \underline{\mathrm{r}}_{\mathrm{Black}}\left(t+\upsilon\right)$ (i.e. the Black (1995) definition for the diffusion process of the shadow short rate subject to the ZLB constraint) and $\mathrm{A}(t+\tau)=1$ (i.e. the terminal payoff for a bond). The resulting expression for $\underline{\mathrm{P}}_{\mathrm{Black}}\left(t,\tau\right)$ is therefore:

$$\underline{\mathrm{P}}_{\mathrm{Black}}\left(t,\tau\right) = \mathbb{E}_t^{\mathbb{Q}}\left\{\exp\left(-\int_0^\tau \underline{\mathrm{r}}_{\mathrm{Black}}\left(t+\upsilon\right)\mathrm{d}\upsilon\right)\right\} \tag{18}$$

Forward rates in the Black (1995) framework are defined with the following standard term structure definition, i.e.:

$$\underline{\mathrm{f}}_{\mathrm{Black}}\left(t,\tau\right) = -\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\underline{\mathrm{P}}_{\mathrm{Black}}\left(t,\tau\right)\right] \tag{19}$$

although this expression is necessarily implicit given the absence of closed-form analytic solutions for $\underline{\mathrm{P}}_{\mathrm{Black}}\left(t,\tau\right)$.

In the ZLB framework, I indirectly define the entire ZLB forward rate curve via equation 17 using $\check{\mathrm{r}}(t+\upsilon) = \mathrm{r}(t+\upsilon)$ (i.e. the shadow short rate itself) and $\mathrm{A}(t+\tau) = \min\left\{1, \mathrm{P}\left(t+\upsilon,\delta\right)\right\}$ (i.e. the terminal payoff for a multiple-finite-step ZLB bond). That set-up results in the multiple-finite-step ZLB bond price $\underline{\mathrm{P}}(t,\upsilon+\delta)$ as an intermediate step, i.e.:

$$\underline{\mathrm{P}}\left(t,\upsilon+\delta\right) = \mathbb{E}_t^{\mathbb{Q}}\left\{\exp\left(-\int_0^\tau \mathrm{r}\left(\upsilon\right)\mathrm{d}\upsilon\right)\cdot\min\left\{1,\mathrm{P}\left(t+\tau,\delta\right)\right\}\right\} \tag{20}$$

Following the discussion in section 2.3, ZLB forward rates $\underline{\mathrm{f}}(t,\tau)$ associated with $\underline{\mathrm{P}}(t,\tau+\delta)$ are then:

$$\underline{\mathrm{f}}\left(t,\tau\right) = \lim_{\tau\to 0}\left\{-\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\log\left[\underline{\mathrm{P}}\left(t,\tau+\delta\right)\right]\right\} \tag{21}$$

and bond prices subject to the ZLB constraint are:

$$\underline{\mathrm{P}}\left(t,\tau\right) = \exp\left[-\int_0^\tau \underline{\mathrm{f}}\left(t,\upsilon\right)\mathrm{d}\upsilon\right] \tag{22}$$

The ZLB bond price $\underline{\mathrm{P}}(t,\tau)$ could also be defined, in principle, via equation 17 using $\check{\mathrm{r}}(t+\upsilon) = \underline{\mathrm{r}}(t+\upsilon)$ and $\mathrm{A}(t+\tau)=1$, i.e.:

$$\underline{\mathrm{P}}\left(t,\tau\right) = \mathbb{E}_t^{\mathbb{Q}}\left\{\exp\left(-\int_0^\tau \underline{\mathrm{r}}\left(t+\upsilon\right)\mathrm{d}\upsilon\right)\right\} \tag{23}$$

where $\underline{\mathrm{r}}(t+\upsilon)$ is the implicit ZLB diffusion process required to reproduce the prices $\underline{\mathrm{P}}(t,\tau)$ from equation 22. Written in this way, it is clear that the result $\underline{\mathrm{P}}(t,\tau) \neq \underline{\mathrm{P}}_{\mathrm{Black}}\left(t,\tau\right)$ from section 2.4.1 implies that:

$$\mathbb{E}_t^{\mathbb{Q}}\left\{\exp\left(-\int_0^\tau \underline{\mathrm{r}}\left(t+\upsilon\right)\mathrm{d}\upsilon\right)\right\} \neq \mathbb{E}_t^{\mathbb{Q}}\left\{\exp\left(-\int_0^\tau \underline{\mathrm{r}}_{\mathrm{Black}}\left(t+\upsilon\right)\mathrm{d}\upsilon\right)\right\} \tag{24}$$

Hence, the different short rate diffusion processes under the risk-neutral $\mathbb{Q}$ measure are different for the Black (1995) and ZLB frameworks. That difference, in turn, arises from the subtlely but fundamentally different mechanism by which each framework converts shadow short rates into "real world" short rates that respect the ZLB. The difference also impacts on the nature of the term structure associated with each framework.

In summary, the Black (1995) framework begins with a simple and explicit process for the risk-neutral short rate subject to the ZLB, but results in a relatively complex and implicit term structure. Conversely, the ZLB framework effectively defines a complex and implicit process for the risk-neutral ZLB short rate, but results in a relatively simple and explicit term structure.

To conclude this section, I introduce the natural follow-up question to the discussion above, i.e.: which framework is "best"? To briefly respond, I believe the answer is an open issue at present. The ZLB framework certainly has an advantage over the Black (1995) framework from the perspective of practical implementation, as sections 3 and 4 will illustrate. However, the comparison from a theoretical perspective is unresolved because both frameworks are essentially based on mechanical/statistical definitions for the short rate subject to the ZLB rather than an explicit theoretical foundation. To be clear on this point, GATSMs themselves have long had explicit theoretical foundations to their statistical specifications,[15] but I am not aware of such foundations for either the Black (1995) framework or my proposed ZLB framework (although, heuristically, both frameworks certainly appear "sensible"). Developing theoretical justifications for the Black (1995) and ZLB frameworks and comparing them is well beyond the scope of this article, but as I discuss in section 5, will be an important topic for future work. Similarly, as also discussed in section 5, a direct empirical comparison may at least resolve the less ambitious question of which framework better represents the data.

3 The ZLB-GATSM framework

In this section, I develop the ZLB-GATSM framework using the general ZLB framework outlined in section 2.3. To establish notation, section 3.1 outlines the generic GATSM specification that I use to represent the shadow-GATSM term structure. While GATSM summaries are available in many articles and textbooks, I generally use Chen (1995) as a convenient point of reference because it contains the explicit multifactor expressions for GATSM bond and option prices that I use respectively in sections 3.2 and 3.3. However, I also refer to Dai and Singleton (2002) pp. 437-8 for the compactness of its matrix notation in some instances.

Section 3.2 derives the shadow-GATSM forward rate $\mathrm{f}(t,\tau)$ and section 3.3 derives the ZLB-GATSM option effect $\mathrm{z}(t,\tau)$.[16] In section 3.4 I combine the results from

[15] For example, Cox, Ingersoll, and Ross (1985a) originally provided a general equilibrium basis for term structure models (including the one-factor GATSM) based on a representative-agent economy, and that approach has since been extended by many authors to provide theoretical foundations for multifactor GATSMs; see, for example, Berardi and Esposito (1999) and Berardi (2009). More recently, Wu (2006) also shows how GATSMs may be given an explicit foundation within dynamic stochastic general equilibrium models.

[16] Chen (1995) actually provides the specification and results for a two-factor GATSM but notes the ready extension to N factors that I use. 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

sections 3.2 and 3.3 to obtain the closed-form analytic expression for ZLB-GATSM forward rates $\underline{\mathrm{f}}(t,\tau)$, which therefore defines ZLB-GATSM interest rates $\underline{\mathrm{r}}(t,\tau)$ and bond prices $\underline{\mathrm{P}}(t,\tau)$ for the ZLB term structure. Section 3.5 makes a series of observations about the ZLB-GATSM framework and term structure.

3.1 Shadow-GATSM short rate process

I use the standard generic GATSM specification from Chen (1995) to represent the shadow-GATSM short rate, i.e.:

$$\mathrm{r}\left(t\right)=\sum_{n=1}^{N}s_{n}\left(t\right) \tag{25}$$

where $s_n(t)$ are the N state variables that evolve as a correlated Ornstein-Uhlenbeck process under the physical or $\mathbb{P}$ measure, i.e.:

$$\mathrm{d}s_{n}\left(t\right)=\kappa_{n}\left[\mu_{n}-s_{n}\left(t\right)\right]\mathrm{d}t+\sigma_{n}\mathrm{d}W_{n}\left(t\right) \tag{26}$$

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 $s_n(t)$, $W_n(t)$ are Wiener processes with $\mathrm{d}W_n(t)\sim N(0,1)\mathrm{d}t$, and $\mathbb{E}\left[\mathrm{d}W_m(t),\mathrm{d}W_m(t)\right]=\rho_{mn}\mathrm{d}t$, where ρ_{mn} are correlations $-1\le\rho_{mn}\le 1$.

Mainly for ease of notation, I adopt the Chen (1995) specification of constant market prices of risk for each factor, which I denote γ_n.[17] However, as in Vasicek (1977), I define the market prices of risk γ_n to be outright positive quantities. That definition means that increases in γ_n intuitively add risk premiums $\sigma_n\gamma_n G(\kappa_n,\tau)$ to forward rates, as in the following section, therefore raising interest rates and lowering bond prices.[18]

The diffusion process under the $\mathbb{P}$ measure in conjunction with the market price of risk specification of course allows one to derive a risk-neutral $\mathbb{Q}$-measure diffusion process for the shadow-GATSM short rate that may be used to price shadow-GATSM securities. However, I need not explicitly undertake such an evaluation in this article because the shadow-GATSM bond price and option price expressions that I require for the ZLB-GATSM framework are already available in Chen (1995).

3.2 Shadow-GATSM forward rates

I derive shadow-GATSM forward rates $\mathrm{f}(t,\tau)$ using the closed-form analytic expression for GATSM bond prices from Chen (1995) within the standard term structure definition $\mathrm{f}(t,\tau)=-\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\mathrm{P}(t,\tau)\right]$. Appendix B.1 contains the details, and the resulting

for multifactor GATSMs appears to begin with Langetieg (1980).

[17] However, the specification could readily be extended to the essentially affine market prices of risk from Duffee (2002); i.e. $\Gamma(t)=\gamma_0+\gamma_1 s(t)$ in obvious matrix notation.

[18] The literature, including Chen (1995), often specifies term structure expressions where terms analogous and related to $\sigma_n\gamma_n G(\kappa_n,\tau)$ are subtracted. However, in such specifications the market prices of risk estimates are negative, so the effect is equivalent to my specification. Dai and Singleton (2002) table 2 is an example that obtains negative market price of risk estimates, and Lund (2001) p. 11 notes that Vasicek (1977) "uses the opposite sign for the market price of risk (which he calls q instead of 'our' λ).".

expression for shadow-GATSM forward rates is:

$$\begin{aligned}
\mathrm{f}(t,\tau) &= \sum_{n=1}^{N}\mu_n + [s_n(t) - \mu_n]\exp(-\kappa_n\tau) \\
&\quad + \sum_{n=1}^{N}\sigma_n\gamma_n G(\kappa_n,\tau) \\
&\quad - \frac{1}{2}\mathrm{Tr}[\Theta(\tau)\Psi] \qquad (27)
\end{aligned}$$

where:

$$G(\kappa_n,\tau) = \frac{1}{\kappa_n}[1 - \exp(-\kappa_n\tau)] \qquad (28)$$

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

$$\Theta_{ij}(\tau) = \rho_{ij}\sigma_i\sigma_j \cdot \kappa_i\kappa_j G(\kappa_i,\tau) G(\kappa_j,\tau) \qquad (29)$$

the matrix Ψ is $\Psi_{ij} = \frac{1}{\kappa_i\kappa_j}$, and Tr[·] is the matrix trace operator.

3.3 ZLB-GATSM option effect

I derive the ZLB-GATSM option effect $\mathrm{z}(t,\tau)$ using the closed-form analytic expressions for GATSM bond prices and call option prices from Chen (1995) within equation 13. Appendix B.2 contains the details, and the resulting ZLB-GATSM option effect is:

$$\mathrm{z}(t,\tau) = -\mathrm{f}(t,\tau)\cdot\left(1 - \mathrm{N}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]\right) + \omega(\tau)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^2\right) \qquad (30)$$

where N[·] is the standard univariate cumulative normal distribution function and $\omega(\tau)$ is the annualized instantaneous option volatility, i.e.:

$$\begin{aligned}
\omega(\tau) &= \lim_{\delta\to 0}\left\{\frac{1}{\delta}\Sigma(\tau,\tau+\delta)\right\} \\
&= \sqrt{\sum_{n=1}^{N}\sigma_n^2\cdot G(2\kappa_n,\tau) + 2\sum_{m=1}^{N}\sum_{n=m+1}^{N}\rho_{mn}\sigma_m\sigma_n\cdot G(\kappa_m+\kappa_n,\tau)} \qquad (31)
\end{aligned}$$

3.4 ZLB-GATSM forward rates

Substituting the respective results for $\mathrm{f}(t,\tau)$ and $\mathrm{z}(t,\tau)$ from equations 27 and 30 into equation 12 gives the generic ZLB-GATSM forward rate expression:

$$\begin{aligned}
\underline{\mathrm{f}}(t,\tau) &= \mathrm{f}(t,\tau) + \mathrm{z}(t,\tau) \\
&= \mathrm{f}(t,\tau) - \mathrm{f}(t,\tau)\left(1 - \mathrm{N}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]\right) + \omega(\tau)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^2\right) \\
&= \mathrm{f}(t,\tau)\cdot\mathrm{N}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right] + \omega(\tau)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^2\right) \qquad (32)
\end{aligned}$$

ZLB-GATSM interest rates $\underline{\mathrm{r}}(t,\tau)$ and bond prices $\underline{\mathrm{P}}(t,\tau)$ are obtained using $\underline{\mathrm{f}}(t,\tau)$ within the respective expressions provided in equations 14 and 15.

3.5 Observations on the ZLB-GATSM framework

One advantage of developing the ZLB-GATSM framework based on the generic GATSM is that the associated observations will apply to any particular ZLB-GATSM specification, regardless of the number of factors and factor inter-relationships. As a related point, the first observation is then that, because any GATSM may be used to represent the shadow term structure, the ZLB-GATSM framework obviously preserves the complete flexibility of the GATSM class of models. The ZLB-GATSM framework provides the explicit modification $\mathrm{z}(t,\tau)$ to ensure that the ZLB is respected by the entire ZLB-GATSM forward rate curve at each point in time, no matter how much shadow-GATSM forward rates evolve below zero. Specifically, as $\mathrm{f}(t,\tau)$ decreases to larger negative values, $\mathrm{N}[\mathrm{f}(t,\tau)/\omega(\tau)]$ and $\exp\left(-\frac{1}{2}[\mathrm{f}(t,\tau)/\omega(\tau)]^2\right)$ both approach zero, so $\lim_{\mathrm{f}(t,\tau)/\omega(\tau)\to-\infty}\underline{\mathrm{f}}(t,\tau)=0$.

The second observation is that the ZLB-GATSM forward rate curve will always be a simple closed-form analytic expression. That is evident from the generic ZLB-GATSM being itself composed of simple closed-form analytic expressions, i.e.: (1) $\mathrm{f}(t,\tau)$, 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 $\mathrm{N}[\mathrm{f}(t,\tau)/\omega(\tau)]$, which is a standard function that is well tabulated or readily approximated with closed-form analytic expressions; and (4) the scalar exponential function $\exp\left(-\frac{1}{2}[\mathrm{f}(t,\tau)/\omega(\tau)]^2\right)$.

Third, ZLB-GATSM interest rates $\underline{\mathrm{r}}(t,\tau)$ for any given time to maturity τ must be obtained via numerical integration over horizon/time to maturity, a property that arises from the Gaussian context.[19] However, the complexity of such integrals remain invariant to the specification of the ZLB-GATSM because, as already noted earlier, the ZLB-GATSM forward rate curve will always be a simple closed-form analytic expression. ZLB-GATSM bond prices $\underline{\mathrm{P}}(t,\tau)$ for any given time to maturity τ are just the scalar exponential of the same univariate numerical integral negated.

Fourth, the relatively straightforward evaluation of ZLB-GATSM interest rates and bond prices along with multivariate-normal transition densities for the ZLB-GATSM state variables means that ZLB-GATSMs will retain a large degree of tractability for empirical applications and estimations.

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

[19] I have tried, obviously without success, to derive closed-form analytic expressions for $\underline{\mathrm{r}}(t,\tau)$, and hence $\underline{\mathrm{P}}(t,\tau)$. The relatively simple form of $\underline{\mathrm{f}}(t,\tau)$ 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.

4 Illustrating the ZLB-GATSM framework

In this section, I illustrate the salient features of the ZLB-GATSM framework using a two-factor ZLB-GATSM based on the two-factor GATSM from Krippner (2012). What I will hereafter call the ZLB-GATSM(2), to be outlined in section 4.1, is ideal for the purposes of demonstration because: (1) it is parsimonious enough to facilitate simple and transparent examples; (2) it provides sufficient flexibility to produce plausible term structures (and emphasize that the ZLB-GATSM framework readily accommodates multifactor shadow-GATSM models); (3) it contains unit-root dynamics that result in negative shadow forward rates and interest rates at very long horizons, thereby providing for an illustration of how that issue is resolved within the ZLB-GATSM framework; and (4) parameter estimates for the shadow-ZLB-GATSM(2) may be taken directly from Krippner (2012). The latter avoids the complexities of a complete empirical estimation that, while obviously a desirable extension, would be beyond the scope of the present article for the reasons discussed in section 5.

Section 4.2 focuses on the cross-sectional perspective of the ZLB-GATSM(2). The examples illustrate how, at a given point in time, the ZLB-GATSM(2) transforms the shadow-GATSM term structure into ZLB-GATSM term structure. Section 4.3 focuses on the time-series perspective of the ZLB-GATSM(2), showing the evolution of the estimated state variables and the associated shadow short rate obtained from United States (U.S.) term structure data.

4.1 The ZLB-GATSM(2)

The shadow-GATSM(2) that underlies the ZLB-GATSM(2) is the two-factor GATSM from Krippner (2012). The shadow-GATSM(2) is very parsimonious because it essentially specifies long-run levels of zero for the two state variables and a zero rate of mean reversion for the first factor. Appendix C details how the Krippner (2012) model may be replicated within the generic shadow-GATSM specification from section 3.2. The resulting shadow-GATSM(2) forward rate expression, with the non-zero rate of mean reversion for the second factor set to λ for notational convenience, is:

$$\begin{aligned} \mathrm{f}(t,\tau) &= s_1(t) + s_2(t)\cdot\exp(-\lambda\tau) \\ &\quad +\sigma_1\gamma_1\cdot\tau + \sigma_2\gamma_2\cdot G(\lambda,\tau) \\ &\quad -\sigma_1^2\cdot\frac{1}{2}\tau^2 - \sigma_2^2\cdot\frac{1}{2}[G(\lambda,\tau)]^2 - \rho\sigma_1\sigma_2\cdot\tau G(\lambda,\tau) \end{aligned} \tag{33}$$

and the ZLB-GATSM(2) annualized instantaneous option volatility is:

$$\omega(\tau) = \sqrt{\sigma_1^2\cdot\tau + \sigma_2^2\cdot G(2\lambda,\tau) + 2\rho\sigma_1\sigma_2\cdot\tau G(\lambda,\tau)} \tag{34}$$

To briefly provide the intuition of the ZLB-GATSM(2): $s_1(t)$ is the "Level" component of the shadow forward rate, and innovations to $s_1(t)$ shift the shadow forward rate curve equally at all horizons/times to maturity); and $s_2(t)\cdot\exp(-\lambda\tau)$ is the "Slope" component, and innovations to $s_1(t)$ shift short-horizon/time to maturity shadow forward rates by more than long-horizon/time to maturity shadow forward rates.[20] Following the discussion in Krippner (2012) pp. 12-14, the first line of equation 33 represents the expected path of the shadow short rate (as at time t and as a function of horizon τ) which I hereafter denote concisely as $\mathbb{E}_t\left[\mathrm{r}\left(t+\tau\right)\right]$, the second line represents risk premiums due to the combined effect of the quantities and market prices of risk (i.e. innovation volatilities and the compensation required by the market to accept the unanticipated price effects associated with those innovations), and the third line represents the volatility effect in the shadow forward rate which captures the influence of volatility on expected returns due to Jensen's inequality (i.e. 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).

The parameters for the shadow-GATSM(2) are the point estimates of the two-factor GATSM from Krippner (2012), p. 22, table 3, i.e.: $\lambda = 0.3884$, $\gamma_1 = 0.1435$, $\gamma_2 = 0.2895$, $\sigma_1 = 0.0172$, $\sigma_2 = 0.0250$, and $\rho = 0.4098$.[21]

ZLB-GATSM(2) forward rates are therefore defined by equations 33 and 34, the parameter set $\mathbb{A} = \{\lambda, \gamma_1, \gamma_2, \sigma_1, \sigma_2, \rho\}$, and the ZLB-GATSM expression in equation 32, which I repeat here for convenience:

$$\underline{\mathrm{f}}\left(t,\tau\right) = \mathrm{f}\left(t,\tau\right)\cdot \mathrm{N}\left[\frac{\mathrm{f}\left(t,\tau\right)}{\omega\left(\tau\right)}\right] + \omega\left(\tau\right)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}\left(t,\tau\right)}{\omega\left(\tau\right)}\right]^2\right) \tag{35}$$

ZLB-GATSM(2) interest rates for any given time to maturity τ are $\underline{\mathrm{r}}(t,\tau) = \frac{1}{\tau}\int_0^\tau \underline{\mathrm{f}}(t,\upsilon)\mathrm{d}\upsilon$. I calculate the latter by numerical integration using the rectangular/mid-ordinate rule with constant increments $\Delta\tau$, i.e.:

$$\underline{\mathrm{r}}\left(t,\tau\right) = \frac{1}{\tau}\left[\Delta\tau\sum_{i=1}^{I}\underline{\mathrm{f}}\left(t,\left[i-\frac{1}{2}\right]\Delta\tau\right)\right] \tag{36}$$

Equation 36 conveniently turns into a arithmetic mean, because $I = \frac{\tau}{\Delta\tau}$, so $\Delta\tau = \frac{\tau}{I}$, and then $\frac{1}{\tau}\Delta\tau = \frac{1}{\tau}\frac{\tau}{I} = \frac{1}{I}$. Therefore:

$$\begin{aligned}\underline{\mathrm{r}}\left(t,\tau\right) &= \frac{1}{I}\sum_{i=1}^{I}\underline{\mathrm{f}}\left(t,\left[i-\frac{1}{2}\right]\Delta\tau\right)\\ &= \mathrm{mean}\left\{\underline{\mathrm{f}}\left(t,0.5\Delta\tau\right),\ldots,\underline{\mathrm{f}}\left(t,\tau-0.5\Delta\tau\right)\right\}\end{aligned} \tag{37}$$

I ensure all integrals are accurate to within 5e-8, which is 5/10000th of a basis point.[22]

[20] Diebold and Li (2006) contains further discussion on the intuition of the Level and Slope components from the perspective of the Nelson and Siegel (1987) model, which is interpreted in Diebold and Li (2006) as a non-arbitrage-free latent three-factor (i.e. Level, Slope, and Curvature factors) dynamic term structure model. Christensen, Diebold, and Rudebusch (2011) is the arbitrage-free analogue of Diebold and Li (2006).

[21] The GATSM(2) estimates from Krippner (2012) are estimated from the same dataset to be outlined in section 4.3 of the present article, but over the period 1988 to 2002.

[22] One 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.

For comparison to the ZLB-GATSM(2) interest rates, I also calculate shadow-GATSM(2) interest rates using the following closed-form analytic expression:

$$
\begin{aligned}
\mathrm{r}(t,\tau) &= s_1(t) + s_2(t)\cdot\frac{1}{\tau}G(\lambda,\tau) \\
&\quad +\sigma_1\gamma_1\cdot\frac{1}{2}\tau + \sigma_2\gamma_2\cdot\frac{1}{\lambda}\left[1-\frac{1}{\tau}G(\lambda,\tau)\right] \\
&\quad -\sigma_1^2\cdot\frac{1}{6}\tau^2 - \sigma_2^2\cdot\frac{1}{\lambda^2}\left[\frac{1}{2}+\frac{1}{\tau}G(\lambda,\tau)+\frac{1}{2\tau}G(2\lambda,\tau)\right] \\
&\quad -\rho\sigma_1\sigma_2\cdot\frac{1}{\lambda^2}\left[\frac{1}{2}\lambda\tau-\frac{1}{\tau}G(\lambda,\tau)-\frac{1}{\tau}\exp(-\lambda\tau)\right] \qquad (38)
\end{aligned}
$$

which I derive via $\frac{1}{\tau}\int_0^\tau \mathrm{x}(t,\upsilon)\mathrm{d}\upsilon$ for each component of the shadow-GATSM(2) forward rate expression in equation 33.

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, in this example I set the ZLB-GATSM(2) state variables to $s_1(t) = 0.05$ and $s_1(t) = -0.10$, which gives a shadow short rate value of $\mathrm{r}(t) = s_1(t) + s_2(t) = -0.05$, or -5 percent.

As indicated in the top-left subplot, the expected path of the shadow short rate $\mathbb{E}_t^{\mathbb{Q}}[\mathrm{r}(t+\tau)]$ rises from its initial value of -5 percent to five percent (the value of $s_1(t)$, which sets the long-horizon level for $\mathbb{E}_t^{\mathbb{Q}}[\mathrm{r}(t+\tau)]$). Regarding intermediate horizons, $\mathbb{E}_t[\mathrm{r}(t+\tau)]$ remains negative for around 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 $\mathrm{r}(t+\tau)$ as a function of horizon τ.

The middle-left subplot shows the shadow forward rate $\mathrm{f}(t,\tau)$ and the option effect $\mathrm{z}(t,\tau)$. $\mathrm{f}(t,\tau)$ has a similar profile to $\mathbb{E}_t[\mathrm{r}(t+\tau)]$ but differs by the shadow risk premium and volatility effect components. Regarding $\mathrm{z}(t,\tau)$, there are two points of note: (1) $\mathrm{z}(t,\tau)$ takes on its highest value for $\tau = 0$, which in turn reflects the certainty that $\mathrm{r}(t)$ is negative (so the option to hold physical currency is "in the money" and exercised, thereby offsetting the negative return that investors would otherwise face on the shadow term structure in the absence of the option effect from the availability of physical currency); and (2) $\mathrm{z}(t,\tau)$ declines but remains materially positive for all horizons shown. That pattern in turn reflects the falling but still material probabilities of $\mathrm{r}(t+\tau)$ remaining negative over each horizon, as already indicated by the probability densities for the shadow short rate.

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

The middle-right subplot shows the shadow forward rate $\mathrm{f}(t,\tau)$ along with the associated ZLB forward rate $\underline{\mathrm{f}}(t,\tau)$. The important points to note are: (1) $\underline{\mathrm{f}}(t,\tau)$ remains essentially at zero until around 0.5 years (and it never takes on negative values); and (2) $\underline{\mathrm{f}}(t,\tau)$ remains distinctly above $\mathrm{f}(t,\tau)$ for all horizons, with the difference being the option effect $\mathrm{z}(t,\tau)$ already discussed. The bottom-right subplot shows shadow interest rates $\mathrm{r}(t,\tau)$ and ZLB interest rates $\underline{\mathrm{r}}(t,\tau)$, which again have similar but more gradual profiles relative to their forward rate counterparts.

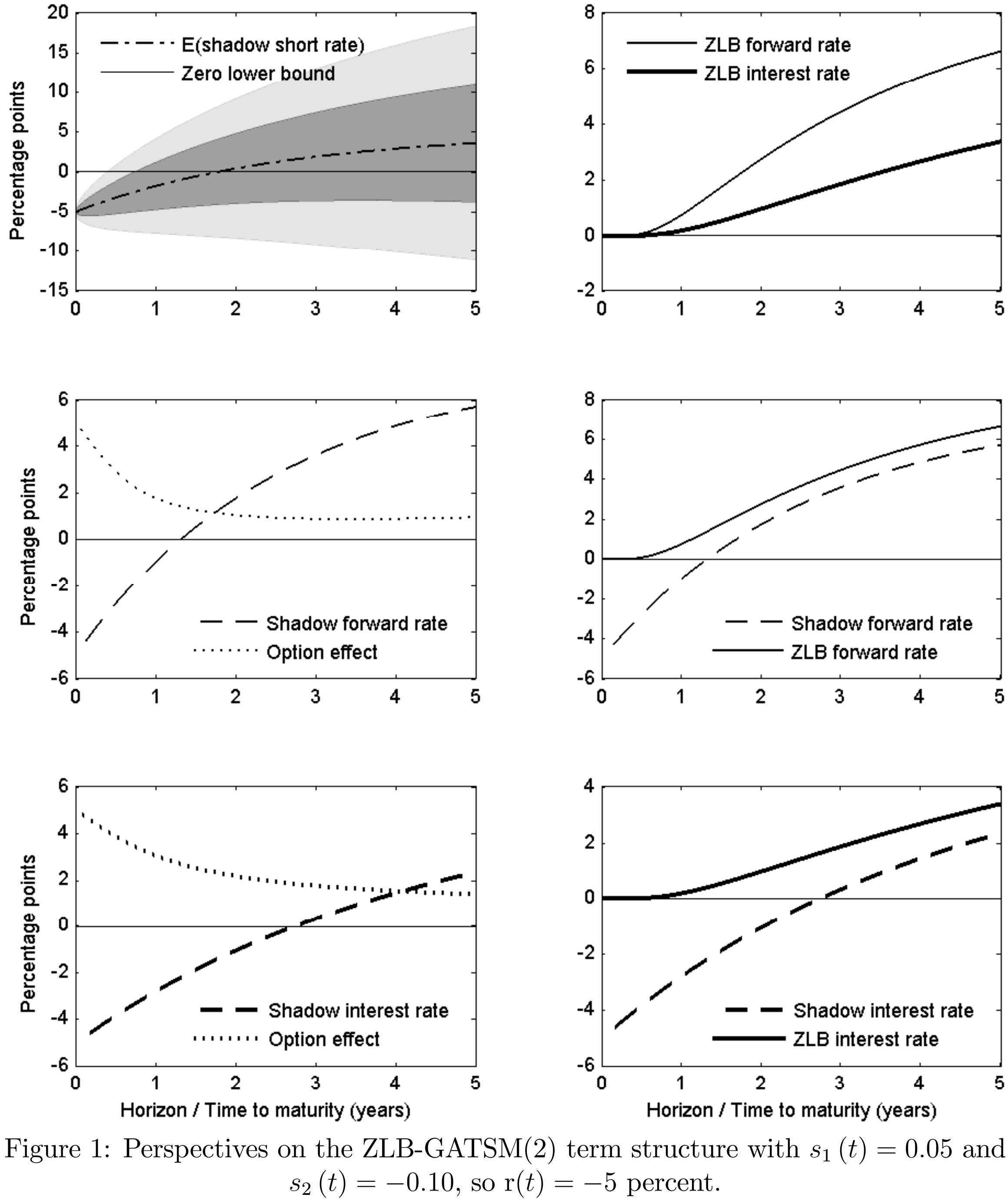

Figure 1: Perspectives on the ZLB-GATSM(2) term structure with $s_1(t) = 0.05$ and $s_2(t) = -0.10$, so $\mathrm{r}(t) = -5$ percent.

Figure 2 shows that the ZLB-GATSM(2) term structure associated with a zero shadow short rate is similar to the shadow-GATSM(2) term structure. In this example, I set the ZLB-GATSM(2) state variables to $s_1(t) = 0.05$ and $s_2(t) = -0.05$. Those values result in a current shadow short rate of $\mathrm{r}(t) = s_1(t)+s_2(t) = 0$, or 0 percent, and $\mathbb{E}_t[\mathrm{r}(t+\tau)]$ again rises to five percent. By comparison to the example from figure 1, forward rates and interest rates are less elevated relative to their shadow counterparts.

That reflects the lower option effect $z(t, \tau)$, which in turn reflects the lower probabilities of $r(t + \tau)$ becoming negative over short and moderate horizons compared to figure 1.

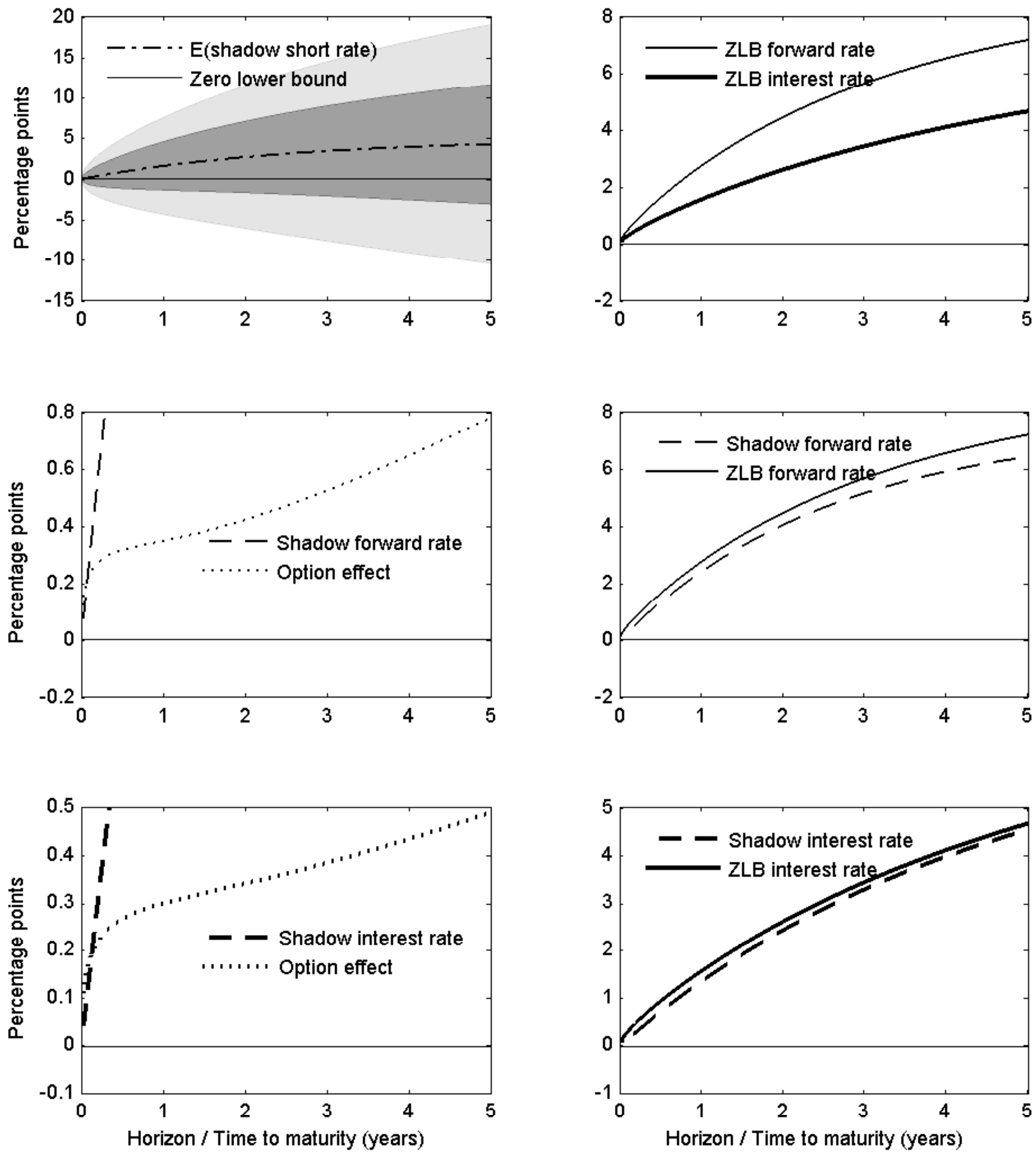

Figure 2: Perspectives on the ZLB-GATSM(2) term structure with $s_1(t) = 0.05$ and $s_2(t) = -0.05$, so $r(t) = 0$ percent.

Figure 3 shows that the ZLB-GATSM(2) term structure associated with a positive shadow short rate is almost identical to the shadow-GATSM(2) term structure for short horizons. In this example, I set the ZLB-GATSM(2) state variables to $s_1(t) = 0.05$ and $s_2(t) = 0.00$. Those value results in a current shadow short rate of $r(t) = s_1(t)+s_2(t) = 0$, or 0 percent, and $\mathbb{E}_t[r(t + \tau)]$ remains constant at five percent. By comparison to the example from figure 2, forward rates and interest rates differ much less from their shadow counterparts for horizons/times to maturity out to around one year. The latter reflects a very low option effect $z(t, \tau)$ for short horizons, which in turn reflects very low probabilities of $\mathrm{r}(t+\tau)$ becoming negative over short horizons. However, figures 2 and 3 become similar from horizons/times to maturity beyond around three years. That reflects the identical Level component for the two examples and the dominance of that component (relative to the Slope component) for moderate horizons and beyond.

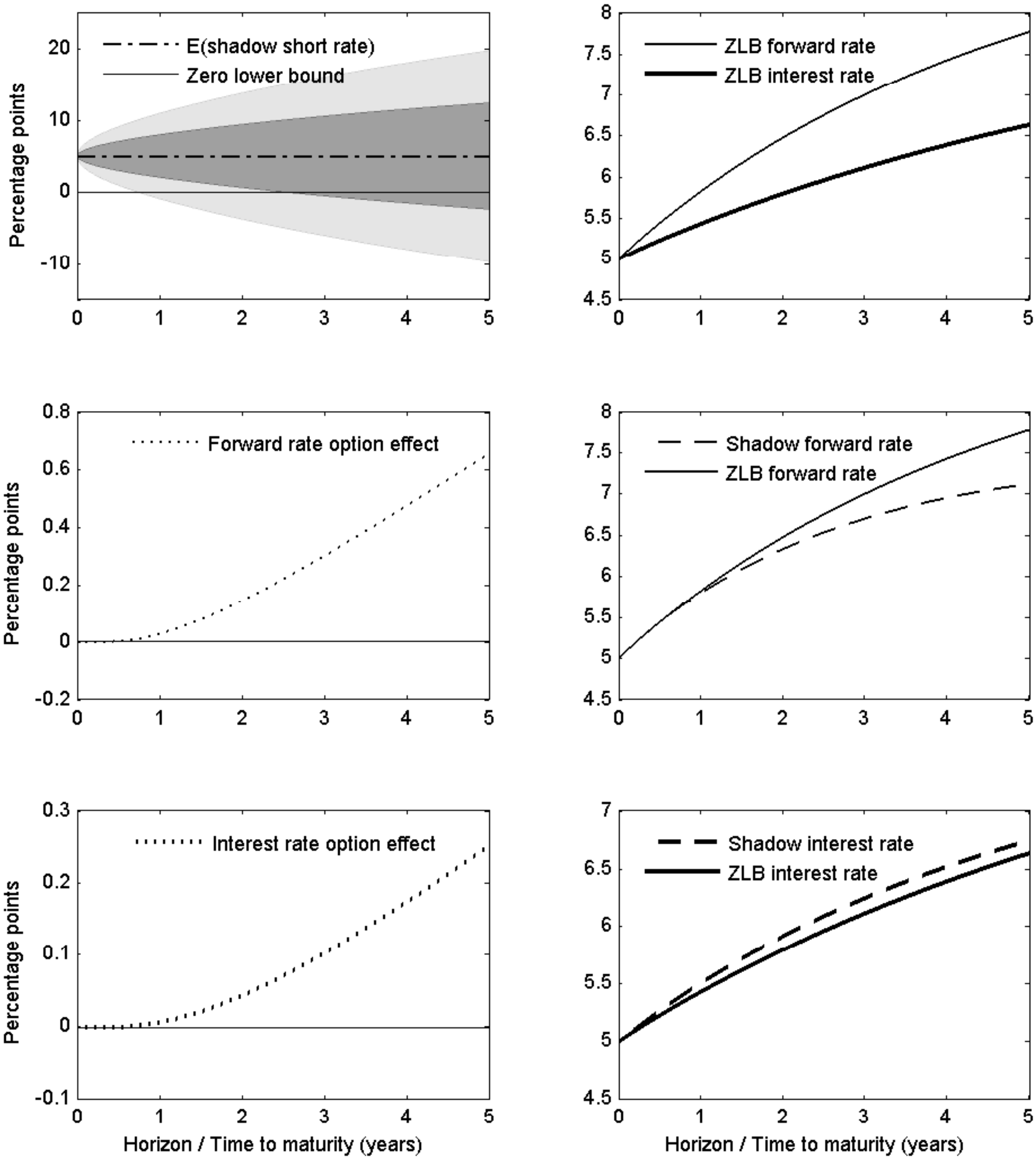

Figure 3: Perspectives on the ZLB-GATSM(2) term structure with $s_1(t) = 0.05$ and $s_2(t) = 0.00$, so $\mathrm{r}(t) = 5$ percent.

As somewhat of an aside to the main point of this section, and mainly out of curiosity, it is worth noting that the ZLB-GATSM framework also accommodates negative shadow forward rates and interest rates that can arise for very long maturities when the shadow-GATSM has very persistent (i.e. slowly mean-reverting) state variables. The shadow-GATSM(2) is an example of such a persistent shadow-GATSM because it contains a unit root process for the Level state variable $s_1(t)$ (which was obtained as the limit of zero mean reversion for the first factor, or $\kappa_1 \to 0$, in appendix C). Appendix D therefore provides a further example of the ZLB-GATSM(2) from a cross-sectional perspective, with the summary results being that the negative shadow forward rates and interest rates for very long horizons/times to maturity are transformed into non-negative ZLB forward rates and interest rates.

4.3 Dynamic/time-series perspective

This section demonstrates the time-series perspective of the ZLB-GATSM framework with an illustrative estimation. Specifically, I use the ZLB-GATSM(2) with calibrated model parameters as specified in section 4.1 to estimate the state variables $s_1(t)$ and $s_2(t)$ from U.S. term structure data. I then discuss the resulting shadow short rate series $\mathrm{r}(t) = s_1(t) + s_2(t)$ from the perspective of the evolving stance of U.S. monetary policy.

The term structure data I use have the same maturities as Krippner (2012), specifically the end of month 3- and 6-month U.S. Treasury bill rates (from the Federal Reserve Economic Database [FRED] 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 (2007) and maintained on the Federal Reserve Board Research Data website. The sample starts at December 1986 and finishes at January 2012 (the last observation at the time of writing).

With calibrated parameters, the estimation of the ZLB-GATSM(2) state variables may simply be undertaken via non-linear least squares at each point in time. Specifically, the objective is to optimize the following system:

$$\begin{bmatrix} \underline{\mathrm{r}}(t, \tau_1) \\ \vdots \\ \underline{\mathrm{r}}(t, \tau_K) \end{bmatrix} = \begin{bmatrix} \underline{\mathrm{r}}(\{s_1(t), s_2(t)\}, \mathbb{A}, \tau_1) \\ \vdots \\ \underline{\mathrm{r}}(\{s_1(t), s_2(t)\}, \mathbb{A}, \tau_K) \end{bmatrix} + \begin{bmatrix} \varepsilon(t, \tau_1) \\ \vdots \\ \varepsilon(t, \tau_K) \end{bmatrix} \tag{39a}$$

$$\text{minimize} \quad : \quad \mathrm{SSR} = \sum_{k=1}^{K} [\varepsilon(t, \tau_k)]^2 \tag{39b}$$

where $\underline{\mathrm{r}}(t, \tau_k)$ are the interest rate data at time t for time to maturity τ_k; the expressions $\underline{\mathrm{r}}(\{s_1(t), s_2(t)\}, \mathbb{A}, \tau_k)$ represent the estimated ZLB-GATSM(2) interest rates as a function of the estimated state variables $s_1(t)$ and $s_2(t)$ at time t, the parameter set $\mathbb{A}$ noted in section 4.1, and the time to maturity τ_k; $\varepsilon(t, \tau_k)$ are the residuals; and k runs from 1 to K, indexing each of the data points being used to represent the term structure at time t (so $K = 12$ in the present application).

As an example of the estimation, figure 4 illustrates the data and results for the last observation of term structure data in the sample. While a comprehensive empirical assessment is not the focus of this article, it is worth noting that the fit to the data is good, indeed, surprisingly so because the ZLB-GATSM(2) uses just two state variables and the parameters are calibrated from a model estimated a decade prior to 2012 (i.e. the GATSM(2) from Krippner (2012) is estimated over the period from 1988 to 2002). Nevertheless, the negative value for $s_1(t)$ implies a long-horizon expectation for negative short rates, and that puzzling result from an economic perspective should lead one to be cautious about interpreting the results from the illustrative estimation too literally. The issue likely lies with the risk premium calibrations, as discussed in the context of the following paragraph.

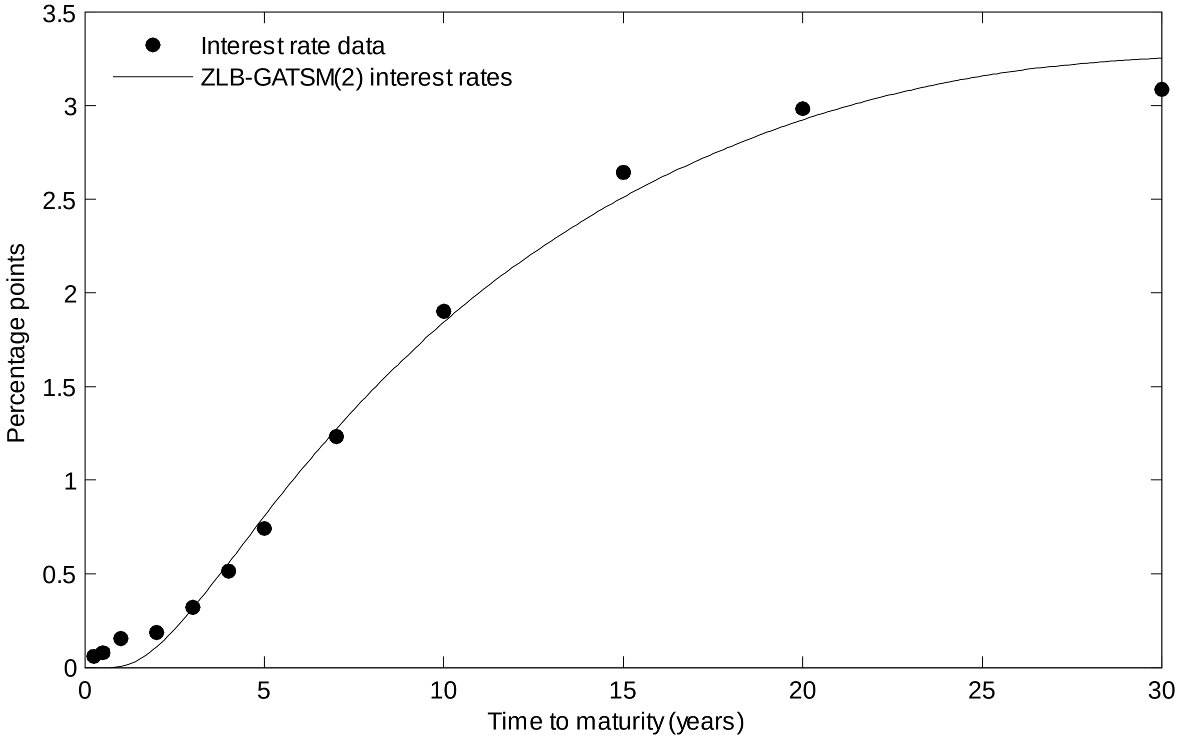

Figure 4: January 2012 interest rate data and the estimated ZLB-GATSM(2) interest rate curve. The Level and Slope state variables are respectively $s_1(t) = -3.61$ percent and $s_2(t) = -3.59$ percent, and the estimated shadow short rate is $\mathrm{r}(t) = -7.19$ percent.

Figure 5 shows the estimated time series for the ZLB-GATSM(2) Level state variable $s_1(t)$ and, for comparison, the 30-year interest rate data. The profiles of the two series match closely over the entire sample, as would be expected because the 30-year interest rate data should reflect long-horizon expectations of the short rate and the ZLB-GATSM(2) Level state variable $s_1(t)$ is the model component representing such expectations. The relatively constant difference between the two series mainly reflects that risk premiums are embedded in the 30-year interest rate data, but not in $s_1(t)$. That is, $\mathbb{E}_t[\mathrm{r}(t+\tau)] = s_1(t) + s_2(t) \cdot \exp(-\lambda\tau)$ does not include a risk premium component, which is explicitly captured elsewhere in the shadow forward rate expression as $\sigma_1\gamma_1 \cdot \tau + \sigma_2\gamma_2 \cdot G(\lambda, \tau)$. Note that the negative values for $s_1(t)$ suggest that the calibrated risk premiums, particularly for the Level component of the term structure, may be too large for the data post 2002.

Figure 6 shows the estimated time series for the ZLB-GATSM(2) Slope state variable $s_2(t)$ and, for comparison, the 3-month less 30-year interest rate spread. The profiles usually match fairly closely over the sample, again with a relatively constant difference due mainly to risk premiums. The very evident exception occurs beyond around late 2008, which I discuss below from the perspectives of figures 7 and 8.

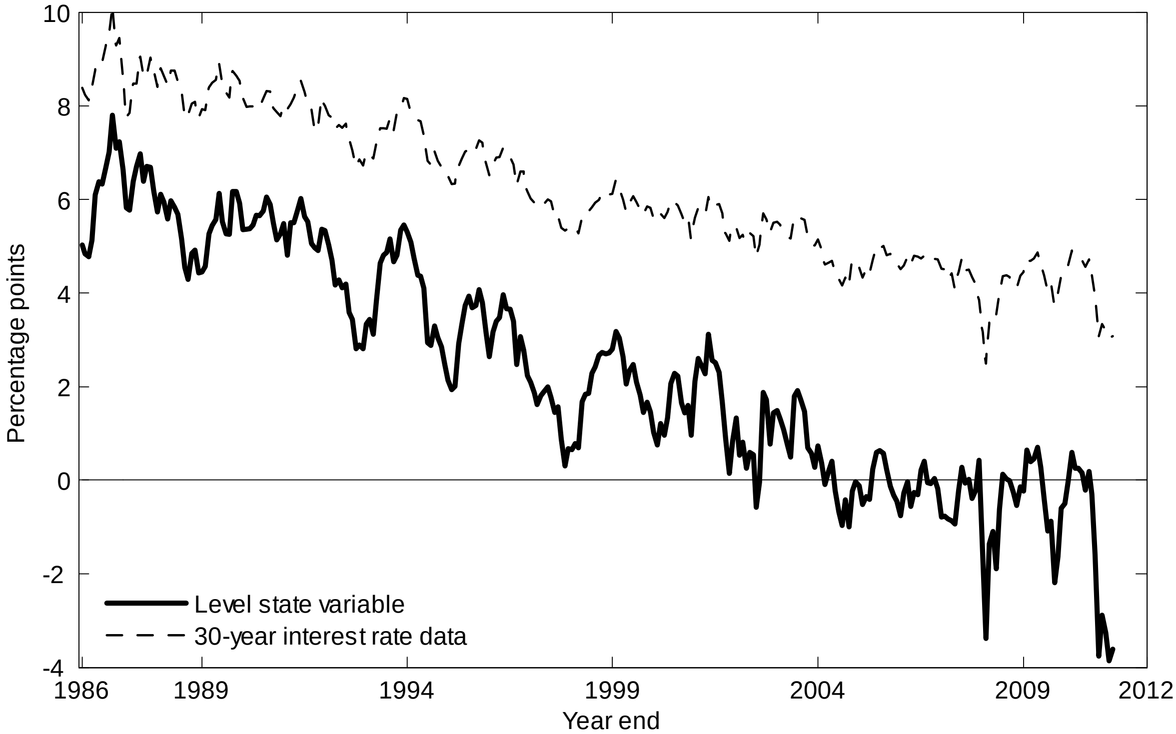

Figure 5: The 30-year interest rate data and estimates of the ZLB-GATSM(2) Level state variable $s_1(t)$.

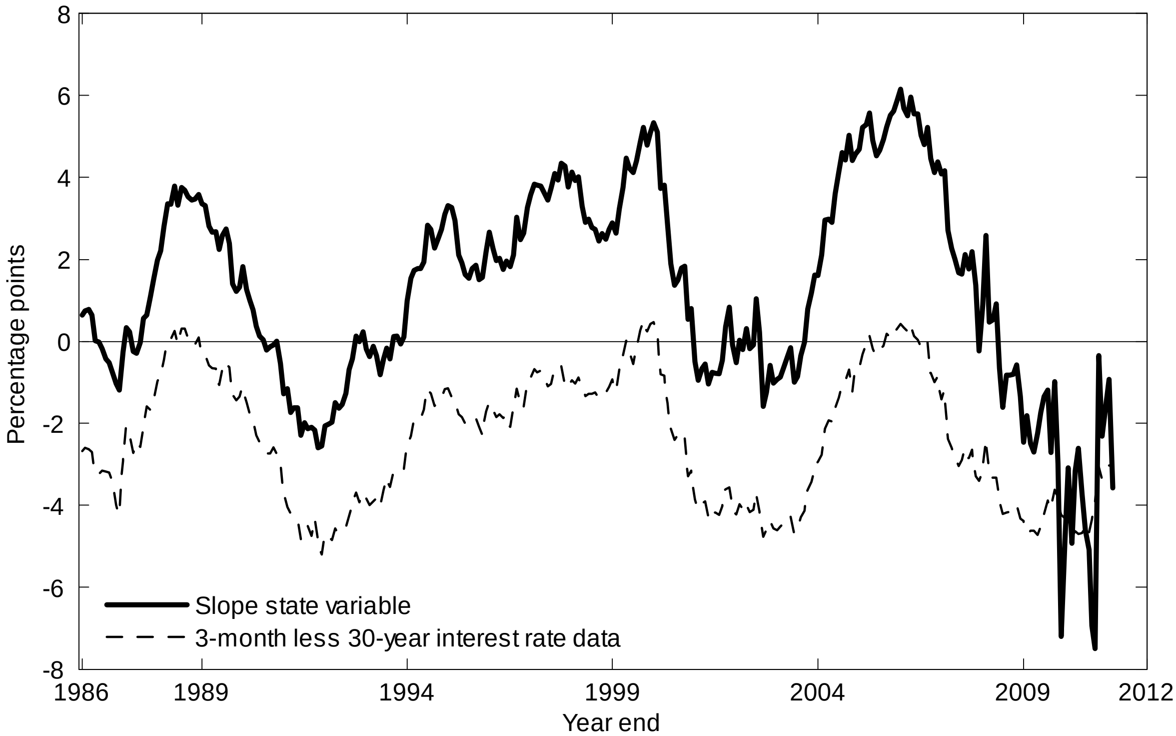

Figure 6: The 3-month less 30-year interest rate spread and estimates of the ZLB-GATSM(2) Slope state variable $s_2(t)$.

Figures 7 and 8 show the 3-month interest rate data and the time series for the model-implied shadow short rate $\mathrm{r}(t) = s_1(t) + s_2(t)$. The first point to note from figure 7 is that r(t) is usually similar to the 3-month interest rate, as would be expected when the term structure data are 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 period of negative values of r(t) occurs from late 2008, which corresponds with events following the Global Financial Crisis (GFC) that began during 2007. The shorter time span of figure 8 provides a better perspective on the downward evolution of the shadow short rate from the start of the GFC until the present.

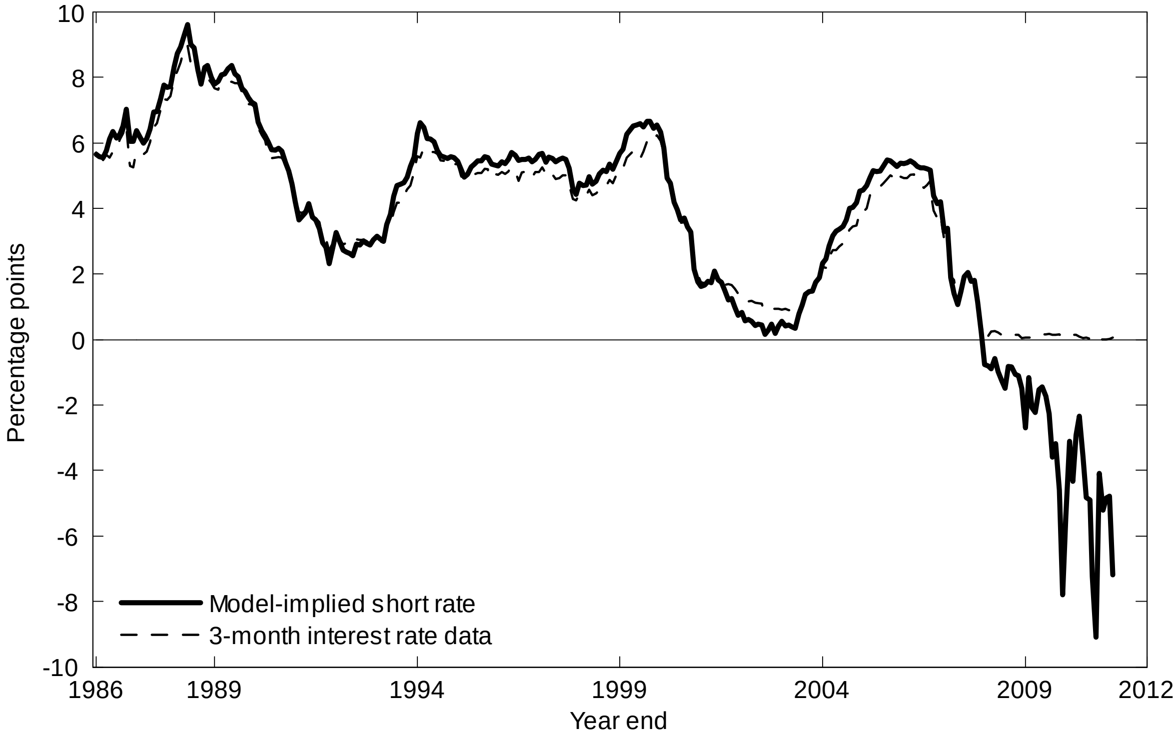

Figure 7: The 3-month interest rate data and estimates of the ZLB-GATSM(2) shadow short rate $\mathrm{r}(t) = s_1(t) + s_2(t)$.

Figure 8 first indicates that the shadow short rate r(t) becomes negative in November 2008. Second, downward movements in r(t) after it becomes negative are broadly consistent with unconventional monetary policy easing events, and the market's anticipation of those events, that I have summarized in the text below figure 8. For example, r(t) continues to decline to more negative values following the FOMC's announcement of "QE1" in November 2008 (the first program of "quantitative easing" [as labeled by the market] involving Federal Reserve purchases of \$1,700 billion of mainly mortgage-backed agency securities, and other measures). In the second half of 2010, r(t) drops substantially following an FOMC warning about a slowing economic recovery, and market anticipation of "QE2" (the second program of "quantitative easing", involving Federal Reserve purchases of \$600 billion of U.S. Treasury securities, subsequently announced in November 2010). Similarly, r(t) drops substantially again during the second half of 2011 following 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 FOMC announcement of a maturity extension program for the Federal Reserve's balance sheet via a switch of $400 billion from short-maturity to longer-maturity Treasury securities. Finally, the January 2012 drop in r(t) coincides with the FOMC's projection of the federal funds rate remaining exceptionally low until late-2014, and also a hint of "QE3" should that prove to be necessary.

Regarding upward movements in the shadow short rate, the sharp rise in r(t) from late-2010 to early-2011 coincides with a run of encouraging economic data at that time, which led markets to bring forward expectations of returning to a more "normal" economic environment and associated monetary policy settings. Indeed, at the time of the local peak of the shadow short rate in March 2010, the March 2012 federal funds futures contract implied a market expectation that the federal funds rate would rise to around 2.00 percent over the following two years (i.e. by March 2012).

The third point from figure 8 is that movements in the shadow short rate appear to become more volatile after it becomes negative. That increase might be a genuine phenomenon, perhaps reflecting that a "conventional" positive policy rate is more transparent to market participants and easier to target by the Federal Reserve. Conversely, movements in the shadow short rate after it becomes negative are associated with the lower transparency and uncertain effects of unconventional monetary policy. Alternatively, the volatility increase might simply be an artifact of the estimated shadow short rate being influenced more by longer-maturity interest rates, which tend to be more volatile, once it becomes negative. In any case, for this illustrative estimation, the mean value of -4.97 percent since August 2010 may be a better broad indication of the average stance of monetary policy maintained since the announcement effect of "QE2" rather than literally interpreting the values from any given month.

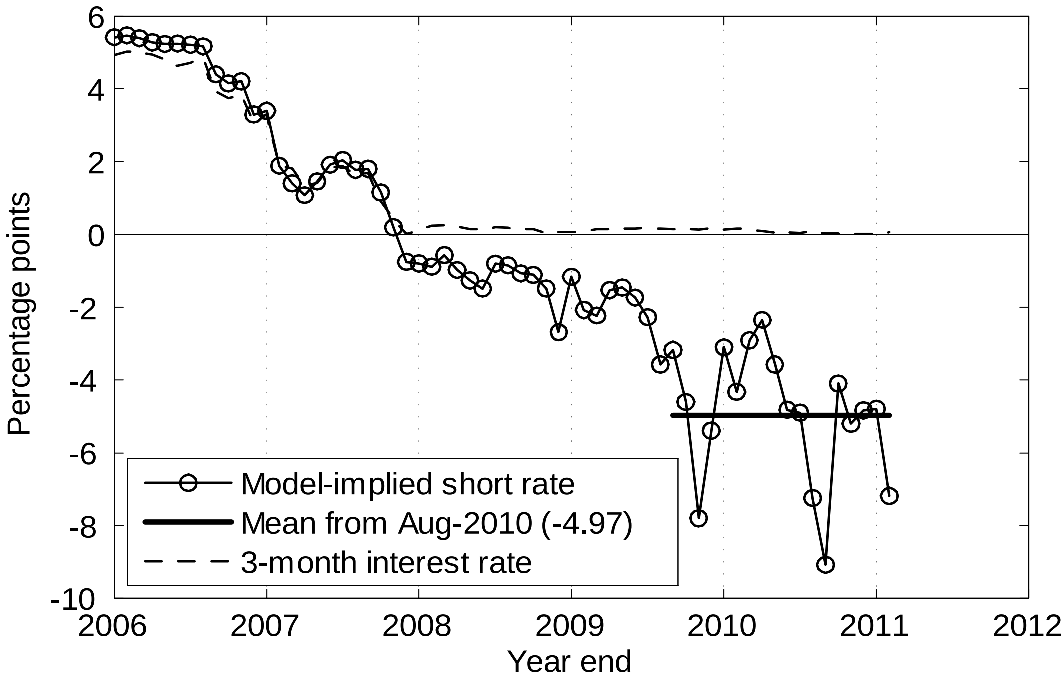

Figure 8: Estimates of the shadow short rate $\mathrm{r}(t) = s_1(t) + s_2(t)$ for the ZLB-GATSM(2) over the period December 2006 to January 2012. The associated monetary policy events for each year are summarized below.

A summary of monetary policy events for the years in figure 8. The dates and the described events and/or quotes are associated with scheduled FOMC meetings unless otherwise noted, and FFR abbreviates "federal funds rate".

- 2007: August 7 FFR held at 5.25 percent; August 10 (unscheduled) FOMC announces liquidity provisions in the first of many such unscheduled events (including term auctions, discount rate cuts, central bank swap arrangements) over 2007 and subsequent years; September 18 FFR cut to 4.75 percent; October 31 FFR cut to 4.50 percent, December 11 FFR cut to 4.25 percent.

- 2008: January 22 (unscheduled) FFR cut to 3.50 percent; January 30 FFR cut to 3.00 percent; March 18 FFR cut to 2.25 percent; April 30 FFR cut to 2.00 percent; (FFR held at 2.00 percent for June 25, August 5, and September 16); September 15 (market event) Lehman bankruptcy; October 8 (unscheduled) FFR cut to 1.50 percent; October 29 FFR cut to 1.00 percent; November 25 (unscheduled) "QE1" announced; December 16 FFR cut to 0-0.25 percent range.

- 2009: no specific easing events but ongoing maintenance of liquidity provisions.

- 2010: August 10 "*recovery ... has slowed*"; August 27 FOMC Chairman Bernanke, at Jackson Hole, foreshadows "QE2"; November 3 (unscheduled) QE2 announced.

- 2011: June 22 "*recovery is continuing ... though somewhat more slowly than the Committee had expected.*"; August 9 "*The Committee currently anticipates ... exceptionally low levels for the federal funds rate at least through mid-2013*"; August 26: FOMC Chairman Bernanke, at Jackson Hole, announces that 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*"; September 21 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.*"

- 2012: January 25 "*economic conditions ... are likely to warrant exceptionally low levels for the federal funds rate at least through late 2014*" and "*The Committee will regularly review the size and composition of its securities holdings and is prepared to adjust those holdings as appropriate to promote a stronger economic recovery in a context of price stability.*"

To summarize figures 7 and 8, 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 (figure 7, pre-2007), and the shadow short rate also reflects unconventional monetary policy easings by taking on increasingly negative values when the term structure data becomes constrained by the ZLB (figure 8, post-September 2008). The latter movements indicate that a potentially valuable application of the ZLB-GATSM framework should be for routinely distilling a quantitative indicator of the stance of monetary policy from term structure data when interest rates are materially constrained by the ZLB.

To complete this section, it is worthwhile briefly comparing the ZLB-GATSM(2) results to those from the existing literature on the ZLB. Unfortunately, no direct comparisons are possible, in terms of a similar model applied to similar data over a similar period. The closest example is Bomfim (2003), which applies a Black-GATSM(2) to U.S. bank-risk data (i.e. London interbank offer rates, or LIBOR, plus associated interest rate swaps) over the period from January 1989 to 17 January 2003. As at that last observation, Bomfim (2003) finds that the GATSM(2) and Black-GATSM(2) term structures are almost identical to each other, both with a short rate of around 1.60 percentage points. My GATSM(2) and ZLB-GATSM(2) term structures as at end-January 2003 are also nearly identical, but with a short rate of 0.60 percentage points. Overall then, the results are broadly similar, allowing for the difference in bank-risk versus government risk data, the model specifications, and the estimation periods.

An even less-direct comparison of the ZLB-GATSM(2) results is to the recent and rapidly growing branch of literature investigating unconventional monetary policy at the ZLB via the movements in longer-maturity interest rates associated with quantities of Large Scale Asset Purchases (LSAPs). Even listing the many articles available is not feasible in the present article, but Williams (2011) p.5, table 1 provides a convenient summary of ten recent quantitative estimates of the effect of a $600 billion LSAP on U.S. Treasury yields, and also a means of converting those effects to an approximate short rate effect.[23] Specifically, the mean and median of the ten point estimates are respectively 0.296 and 0.175 percentage points, and Williams (2011) p. 4 notes that a typical historical response of the 10-year Treasury yield to a 0.75 percentage point cut in the FFR is approximately 0.15-0.20 percentage points. Therefore, scaling the mean and median values up by $2,700/$600 ("QE1" of $1,700 billion, "QE2" of $600 billion, and the maturity extension of $400 billion) and 0.75/0.175 (the mid-point of the FFR/Treasury yield effect) respectively gives approximate estimates of 5.71 and 3.38 percentage points equivalent on the FFR. Of course, those estimates should not be taken too seriously, given my extremely simple method of converting between quantities and the FFR, but from the perspective of this article it is encouraging to see that they are broadly consistent with the ZLB-GATSM(2) results. That is, the LSAP-FFR estimates based on the mean and median span the average ZLB-GATSM(2) shadow short rate of 4.97 percent that has prevailed since August 2010.

In summary, the similarity of the results from two diverse approaches suggest that indicators of the stance of monetary policy from ZLB-GATSMs and LSAP/quantity approaches should complement each other. That is not surprising, because in principle a fundamental relationship should exist between the term structure and quantities/expected quantities of money, albeit measuring either of the latter values has never been easy. Indeed, from that perspective, the ZLB-GATSM may be seen as an approach for directly measuring the interest rate effects of quantity-based monetary policy actions that have been pursued due to the ZLB constraint.

[23] Kozicki, Santor, and Surchnek (2011) provides an more extensive summary of the literature, with table 1 on p. 19 listing 15 articles that quantitatively and/or qualitatively assess the impact of credit easing and LSAPs on financial markets. Table 2 of that article summarizes four assessments of the macroeconomic impact of LSAPs.

5 Extensions and future work

The most immediate extension of the ZLB-GATSM material presented in this article, which I intend to undertake in forthcoming work, is a thorough empirical assessment of the state variables and model parameters of the ZLB-GATSM(2) and/or alternative ZLB-GATSM specifications. Some of the issues to be considered 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-GATSM(2) specification), and whether time-invariant or time-varying market prices of risk are most appropriate. 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 for regime changes suggested by external information. For example, the increase in volatility of the shadow short rate in figure 7 after the FFR reaches the ZLB and the FOMC begin unconventional monetary policy actions might represent a testable regime change.

Once those aspects have been addressed appropriately, I can proceed with a thorough assessment of the empirical properties of the ZLB-GATSM shadow short rate and shadow term structure as indicators of the stance of monetary policy when interest rates are materially constrained by the ZLB. Given that situation exists (at the time of writing) in several major developed economies, relevant data is readily available for such an exercise and the results should prove practically useful to central bank policy makers. Another related practical application of the shadow term structure data obtained from the ZLB-GATSM is to indicate when the market expects the central bank's monetary policy rate of zero to end. Shadow term structure data may also be useful to use in macroeconomic models that assume, implicitly or explicitly, interest rates with Gaussian properties (e.g. DSGE models). For example, transforming ZLB interest data into shadow interest data and using the latter may be more appropriate for some applications, rather than using ZLB data that is known to have non-Gaussian properties near the ZLB.

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 call options to a value above 1 (thereby allowing for slightly negative interest rates). Another degree of flexibility would be to specify separate term structures for borrowing and lending to better represent the effect of central bank floors and ceilings for policy rate settings.

Regarding ZLB term structures in general, a potential extension would be to use non-Gaussian models to represent the shadow term structure within the general ZLB framework, along with their associated option prices. In particular, closed-form analytic expressions for forward rates and option prices are already available for term structure models with independent square-root processes. Hence, models combining the generic GATSM with one or several independent square-root processes might conveniently provide for some time-varying volatility (i.e. heteroskedasticity) while retaining most of the tractability of the 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.

Another extension of the general ZLB framework would be to specify how securities other than bond prices could be priced under the ZLB constraint. For example, the price of ZLB options on ZLB bonds could be specified analogous to equations 17 and 23, although the outcome might not prove as user-friendly as ZLB bond prices themselves.

Finally, as already discussed in section 2.4, further work would be required to resolve (if possible) which of the Black (1995) and ZLB frameworks is "best". Assessing the fit to term structure data and/or forecasting exercises may provide an empirical means of choosing between the two frameworks. However, justifying either (or both) frameworks from a theoretical perspective would arguably provide a more definitive and enduring means of selecting the "best" framework.

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 eliminate negative interest rates by adding to the entire shadow-GATSM forward rate curve an explicit function of maturity that represents the option effect from the present and future availability of physical currency. That option effect is derived using the closed-form analytic expression for GATSM call options on bonds.

Models in the ZLB-GATSM class retain all of the flexibility of the GATSMs on which they are based, and the ZLB-GATSM term structures retain most of the tractability of GATSMs. Specifically, regardless of the specification of the shadow-GATSM (i.e. the number of factors and their inter-relationships), ZLB-GATSM forward rates will always have a simple closed-form analytic expression, involving just exponential functions and the univariate cumulative normal distribution. Those simple expressions mean that numerically integrating ZLB forward rates to obtain ZLB-GATSM interest rates and bond prices will always be elementary.

While I have yet to undertake a full empirical estimation, an illustrative application of a calibrated-parameter two-factor ZLB-GATSM model to U.S. term structure data provides sensible and intuitive results; i.e. the shadow short rate evolves consistently with both conventional and unconventional monetary policy settings; in particular, increasingly negative values of the shadow short rate have coincided with unconventional monetary policy easings undertaken after the U.S. policy rate reached the ZLB in late 2008. The shadow short rate therefore provides a useful gauge of the stance of monetary policy after the U.S. policy rate reached the ZLB in late 2008.

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.

Bomfim, A. (2003). 'interest rates as options': assessing the markets' view of the liquidity trap. *Working Paper, Federal Reserve Board of Governors*.

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 (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.

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 (2007). 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.

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 (2011). A new perspective on Gaussian dynamic term structure models. *Review of Financial Studies 24(3)*, 926–970.

Kim, D. and K. Singleton (2011). Term structure models and the zero bound: an empirical investigation of Japanese yields. *Working paper, Stanford University.*

Kozicki, S., E. Santor, and L. Surchnek (2011). Uncoventional monetary policy: the international experience with central bank asset purchases. *Bank of Canada Review Spring 2011*, 13–25.

Krippner, L. (2012). A theoretical foundation for the Nelson and Siegel class of yield curve models. *Discussion paper, Centre for Applied Macreconomic Analysis 11/2012.*

Langetieg, T. (1980). A multivariate model of the term structure. *Journal of Finance 35(1)*, 71–97.

Lund, J. (2001). Review of continuous-time term-structure models, part I: equilibrium models. *Lecture Notes, Aarhus School of Business.*

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.*

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.

Williams, J. (2011). Unconventional monetary policy: lessons from the past three years. *Federal Reserve Bank of San Francisco Economic Letter 2011-31.*

Wu, T. (2006). Macro factors and the affine term structure of interest rates. *Journal of Money, Credit and Banking 38(7)*, 1847–1875.

A General ZLB forward rate and short rate

Appendix A.1 contains the detailed derivation of the general ZLB forward rate expression $\underline{\mathrm{f}}(t,\tau)$ in equation 12. Appendix A.2 contains the detailed results for the general ZLB short rate $\underline{\mathrm{r}}(t)$ as summarized in equation 16.

A.1 General ZLB forward rate derivation

I begin from the following standard term structure relationship:

$$\begin{aligned}
\underline{\mathrm{f}}(t,\tau) &= -\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\underline{\mathrm{P}}(t,\tau)\right] \\
&= \lim_{\delta\to 0}\left\{-\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\log\left[\underline{\mathrm{P}}(t,\tau+\delta)\right]\right\} \\
\langle\text{Chain rule}\rangle &= \lim_{\delta\to 0}\left\{-\frac{\mathrm{d}}{\mathrm{d}x}\log\left[x\right]\cdot\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\underline{\mathrm{P}}(t,\tau+\delta)\right\} \\
&\quad \left\langle x=\underline{\mathrm{P}}(t,\tau+\delta)\,;\ \frac{\mathrm{d}}{\mathrm{d}x}\log\left[x\right]=\frac{1}{x}\right\rangle \\
&= \lim_{\delta\to 0}\left\{-\frac{1}{\underline{\mathrm{P}}(t,\tau+\delta)}\left[\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\mathrm{P}(t,\tau+\delta)-\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\mathrm{C}(t,\tau,\tau+\delta)\right]\right\} \\
\langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{-\frac{1}{\underline{\mathrm{P}}(t,\tau+\delta)}\right\}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\mathrm{P}(t,\tau+\delta)\right\} \\
&\quad +\lim_{\delta\to 0}\left\{\frac{1}{\underline{\mathrm{P}}(t,\tau+\delta)}\right\}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\mathrm{C}(t,\tau,\tau+\delta)\right\} \qquad (40)
\end{aligned}$$

where "Limit rules" refers to the standard rules of calculus for manipulating limits of analytic functions (in this case the sum of limits,i.e. $\lim_{x\to a}\{f(x)+g(x)\}=\lim_{x\to a}f(x)+\lim_{x\to a}g(x)$ and the product of limits, i.e. $\lim_{x\to a}\{f(x)\cdot g(x)\}=\lim_{x\to a}f(x)\cdot\lim_{x\to a}g(x)$).

The quotient limit rule (i.e. $\lim_{x\to a}\{f(x)/g(x)\}=\lim_{x\to a}f(x)/\lim_{x\to a}g(x)$ if $\lim_{x\to a}g(x)\neq 0$) may be used to evaluate the expression $\lim_{\delta\to 0}\left\{-\frac{1}{\underline{\mathrm{P}}(t,\tau+\delta)}\right\}$ in equation 40, i.e.:

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{1}{\underline{\mathrm{P}}(t,\tau+\delta)}\right\} &= \lim_{\delta\to 0}\left\{\left[\underline{\mathrm{P}}(t,\tau+\delta)\right]^{-1}\right\} \\
\langle\text{Limit rules}\rangle &= \left[\lim_{\delta\to 0}\left\{\mathrm{P}(t,\tau+\delta)-\mathrm{C}(t,\tau,\tau+\delta)\right\}\right]^{-1} \\
\langle\text{Limit rules}\rangle &= \left[\lim_{\delta\to 0}\left\{\mathrm{P}(t,\tau+\delta)\right\}-\lim_{\delta\to 0}\left\{\mathrm{C}(t,\tau,\tau+\delta)\right\}\right]^{-1} \\
&= \left[\mathrm{P}(t,\tau)-0\right]^{-1} \\
&= \frac{1}{\mathrm{P}(t,\tau)} \qquad (41)
\end{aligned}$$

where the result $\lim_{\delta\to 0}\{\mathrm{C}(t,\tau,\tau+\delta)\}=0$ may be obtained taking the limit of a generic call option price expression with a strike price set to 1. For example, from Filipović (2009) p. 109:

$$\mathrm{C}(t,\tau,\tau+\delta)=\mathbb{E}_t^{\mathbb{Q}}\left\{\exp\left(-\int_0^{\tau}\check{\mathrm{r}}(t+\upsilon)\,\mathrm{d}\upsilon\right)\cdot\max\left[\mathrm{P}(t+\tau,\delta)-1,0\right]\right\} \qquad (42)$$

where $\mathbb{E}_t^{\mathbb{Q}}$ is the expectation operator under the risk-neutral $\mathbb{Q}$ measure and $\check{\mathrm{r}}(t+\upsilon)$ is the short rate used to obtain the discount factor for future cashflows. The result of zero follows because the limit of the payoff component in all states is zero, i.e.:

$$\begin{aligned}\lim_{\delta\to 0}\left\{\max\left[\mathrm{P}\left(t+\tau,\delta\right)-1,0\right]\right\} &= \max\left[\lim_{\delta\to 0}\left\{\mathrm{P}\left(t+\tau,\delta\right)-1\right\},0\right] \\ &= \max\left[\mathrm{P}\left(t+\tau,0\right)-1,0\right] \\ &= \max\left[1-1,0\right] \\ &= 0 \end{aligned} \tag{43}$$

The expressions $\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}[\tau+\delta]}\mathrm{P}\left(t,\tau+\delta\right)\right\}$ and $\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}[\tau+\delta]}\mathrm{C}\left(t,\tau,\tau+\delta\right)\right\}$ in equation 40 may respectively be simplified as follows (in both cases beginning with the chain rule):

$$\begin{aligned}\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\mathrm{P}\left(t,\tau+\delta\right)\right\} &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}\left(t,\tau+\delta\right)\cdot\frac{\mathrm{d}\tau}{\mathrm{d}\left[\tau+\delta\right]}\right\} \\ \langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}\left(t,\tau+\delta\right)\right\}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}\tau}{\mathrm{d}\left[\tau+\delta\right]}\right\} \\ \langle\text{Limit rules}\rangle &= \frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}\left(t,\tau\right)\cdot\left[\lim_{\delta\to 0}\left\{\frac{\mathrm{d}\left[\tau+\delta\right]}{\mathrm{d}\tau}\right\}\right]^{-1} \\ &= \frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}\left(t,\tau\right)\cdot\left[\lim_{\delta\to 0}\left\{1\right\}\right]^{-1} \\ &= \frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}\left(t,\tau\right) \end{aligned} \tag{44}$$

$$\begin{aligned}\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\mathrm{C}\left(t,\tau,\tau+\delta\right)\right\} &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{C}\left(t,\tau,\tau+\delta\right)\cdot\frac{\mathrm{d}\delta}{\mathrm{d}\left[\tau+\delta\right]}\right\} \\ \langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{C}\left(t,\tau,\tau+\delta\right)\right\}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}\delta}{\mathrm{d}\left[\tau+\delta\right]}\right\} \\ \langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{C}\left(t,\tau,\tau+\delta\right)\right\}\cdot\left[\lim_{\delta\to 0}\left\{\frac{\mathrm{d}\left[\tau+\delta\right]}{\mathrm{d}\delta}\right\}\right]^{-1} \qquad (45) \\ &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{C}\left(t,\tau,\tau+\delta\right)\right\} \qquad (46) \end{aligned}$$

Substituting the three limit expressions above back into equation 40 gives:

$$\begin{aligned}\underline{\mathrm{f}}\left(t,\tau\right) &= -\frac{1}{\mathrm{P}\left(t,\tau\right)}\left[\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}\left(t,\tau\right)-\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{C}\left(t,\tau,\tau+\delta\right)\right\}\right] \\ &= -\frac{1}{\mathrm{P}\left(t,\tau\right)}\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}\left(t,\tau\right)+\frac{1}{\mathrm{P}\left(t,\tau\right)}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{C}\left(t,\tau,\tau+\delta\right)\right\} \\ &= \mathrm{f}\left(t,\tau\right)+\mathrm{z}\left(t,\tau\right) \end{aligned} \tag{47}$$

where the result $\mathrm{f}(t,\tau)$ arises from re-expressing the definition of forward rates from bond prices, i.e.:

$$\begin{aligned}
\mathrm{f}(t,\tau) &= -\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\mathrm{P}(t,\tau)\right] \\
\langle\text{Chain rule}\rangle &= -\frac{\mathrm{d}}{\mathrm{d}x}\log[x]\cdot\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}(t,\tau) \\
&\qquad \left\langle x=\mathrm{P}(t,\tau)\,;\ \frac{\mathrm{d}}{\mathrm{d}x}\log[x]=\frac{1}{x}\right\rangle \\
&= -\frac{1}{\mathrm{P}(t,\tau)}\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{P}(t,\tau)
\end{aligned} \tag{48}$$

and I have denoted:

$$\mathrm{z}(t,\tau)=\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{C}(t,\tau,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\} \tag{49}$$

A.2 General ZLB short rate derivation

I begin from the following standard term structure relationship:

$$\begin{aligned}
\underline{\mathrm{r}}(t) &= \lim_{\tau\to 0}\left\{\underline{\mathrm{f}}(t,\tau)\right\} \\
&= \lim_{\tau\to 0}\left\{\mathrm{f}(t,\tau)+\mathrm{z}(t,\tau)\right\} \\
\langle\text{Limit rules}\rangle &= \lim_{\tau\to 0}\left\{\mathrm{f}(t,\tau)\right\}+\lim_{\tau\to 0}\left\{\mathrm{z}(t,\tau)\right\} \\
&= \mathrm{f}(t,0)+\lim_{\tau\to 0}\left\{\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{C}(t,\tau,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}\right\} \\
\langle\text{Limit rules}\rangle &= \mathrm{r}(t)+\lim_{\delta\to 0}\left\{\lim_{\tau\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{C}(t,\tau,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}\right\}
\end{aligned} \tag{50}$$

where the results $\lim_{\tau\to 0}\{\mathrm{f}(t,\tau)\}=\mathrm{f}(t,0)=\mathrm{r}(t)$ are term structure definitions for the shadow term structure.

Note that the limit operators for the option effect have been interchanged in the final step above, which is valid because the τ and δ are independent variables. Evaluating the option effect expression begins by interchanging the $\lim_{\tau\to 0}$ and $\frac{\mathrm{d}}{\mathrm{d}\delta}$ operators (again valid due to independence), i.e.:

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\lim_{\tau\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{C}(t,\tau,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}\right\} &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\lim_{\tau\to 0}\left[\frac{\mathrm{C}(t,\tau,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\} \\
\langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\lim_{\tau\to 0}\mathrm{C}(t,\tau,\tau+\delta)}{\lim_{\tau\to 0}\mathrm{P}(t,\tau)}\right]\right\} \\
&= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{C}(t,0,\delta)}{1}\right]\right\}
\end{aligned} \tag{51}$$

The limit of the derivative may be evaluated by substituting the payoff for the expiring option $\mathrm{C}(t,0,\delta)$, i.e.:

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{C}\left(t,0,\delta\right)\right\} &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\max\left\{\mathrm{P}\left(t,\delta\right)-1,0\right\}\right\} \\
&= \lim_{\delta\to 0}\left\{\max\left[\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\mathrm{P}\left(t,\delta\right)-1\right],0\right]\right\} \\
&= \lim_{\delta\to 0}\left\{\max\left[\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{P}\left(t,\delta\right),0\right]\right\} \\
&= \lim_{\delta\to 0}\left\{\max\left[-\mathrm{P}\left(t,\delta\right)\cdot\mathrm{f}\left(t,\delta\right),0\right]\right\} \\
\langle\text{Limit rules}\rangle &= \max\left[\lim_{\delta\to 0}\left\{-\mathrm{P}\left(t,\delta\right)\cdot\mathrm{f}\left(t,\delta\right)\right\},0\right] \qquad (52)
\end{aligned}$$

where the result $\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{P}(t,\delta) = -\mathrm{P}(t,\delta)\cdot\,\mathrm{f}(t,\delta)$ is apparent from re-arranging equation 48 and substituting δ for τ, and the final step uses the limit rule for functions, i.e. $\lim_{x\to a}\left\{g\left[f\left(x\right)\right]\right\} = g\left[\lim_{x\to a}\left\{f\left(x\right)\right\}\right]$. Finally, applying the product limit rule gives:

$$\begin{aligned}
\lim_{\delta\to 0}\left\{-\mathrm{P}\left(t,\delta\right)\cdot\mathrm{f}\left(t,\delta\right)\right\} &= -\lim_{\delta\to 0}\left\{\mathrm{P}\left(t,\delta\right)\right\}\cdot\lim_{\delta\to 0}\left\{\mathrm{f}\left(t,\delta\right)\right\} \\
&= -1\cdot\mathrm{f}\left(t,0\right) \\
&= -\mathrm{r}\left(t\right) \qquad (53)
\end{aligned}$$

Substituting the option effect result back through to the original expression in equation 50 gives the final result:

$$\begin{aligned}
\underline{\mathrm{r}}\left(t\right) &= \mathrm{r}\left(t\right)+\max\left\{-\mathrm{r}\left(t\right),0\right\} \\
&= \max\left\{0,\mathrm{r}\left(t\right)\right\} \qquad (54)
\end{aligned}$$

B ZLB-GATSM forward rate derivation

Appendix B.1 contains the derivations of the shadow forward rate expressions in section 3.2. Appendix B.2 contains the derivations of the option effect expressions in section 3.3.

B.1 Shadow-GATSM forward rates

Chen (1995) p. 349-50 provides the following closed-form analytic expression for GATSM bond prices:[24]

$$\mathrm{P}\left(t,\tau\right) = \exp\left[-H\left(\tau\right)-\sum_{n=1}^{N}s_n\left(t\right)\cdot G\left(\kappa_n,\tau\right)\right] \qquad (55)$$

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

$$G\left(\kappa_n,\tau\right) = \frac{1}{\kappa_n}\left[1-\exp\left(-\kappa_n\tau\right)\right] \qquad (56)$$

[24] Chen (1995) actually uses time t and time of maturity T notation. My time t and time to maturity $\tau = T - t$ notation is equivalent.

and the function $H(\tau)$ for the generic GATSM is most concisely expressed using Dai and Singleton (2002) matrix notation:[25]

$$H(\tau) = -\frac{1}{2}\mathrm{Tr}\left[\tau\Xi(\tau)\Psi\right] + \sum_{n=1}^{N}\left[\mu_n + \frac{\sigma_n\gamma_n}{\kappa_n}\right]\left[\tau - G(\kappa_n,\tau)\right] \tag{57}$$

where Tr[·] is the matrix trace operator, the matrix Ψ is:

$$\Psi_{ij} = \frac{1}{\kappa_i\kappa_j} \tag{58}$$

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

$$\tau\Xi_{ij}(\tau) = \rho_{ij}\sigma_i\sigma_j \cdot \left[\tau - G(\kappa_i,\tau) - G(\kappa_j,\tau) + G(\kappa_i+\kappa_j,\tau)\right] \tag{59}$$

I derive shadow-GATSM forward rates beginning with the standard term structure relationship for forward rates and bond prices, and then substituting the bond price expression from equation 55, i.e.:

$$\begin{aligned}
\mathrm{f}(t,\tau) &= -\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\mathrm{P}(t,\tau)\right] \\
&= -\frac{\mathrm{d}}{\mathrm{d}\tau}\log\left[\exp\left(-H(\tau) - \sum_{n=1}^{N} s_n(t)\cdot G(\kappa_n,\tau)\right)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}\tau}H(\tau) + \sum_{n=1}^{N} s_n(t)\cdot\frac{\mathrm{d}}{\mathrm{d}\tau}G(\kappa_n,\tau)
\end{aligned} \tag{60}$$

The expressions $\frac{\mathrm{d}}{\mathrm{d}\tau}G(\kappa_n,\tau)$ are readily calculated from the definition in equation 56, i.e.:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}G(\kappa_n,\tau) &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{\kappa_n}\left[1-\exp(-\kappa_n\tau)\right]\right) \\
&= \frac{1}{\kappa_n}\cdot\kappa_n\exp(-\kappa_n\tau) \\
&= \exp(-\kappa_n\tau)
\end{aligned} \tag{61}$$

as is the expression $\frac{\mathrm{d}}{\mathrm{d}\tau}H(\tau)$ from the definition in equation 57, i.e.:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}H(\tau) &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(-\frac{1}{2}\mathrm{Tr}\left[\tau\Xi(\tau)\Psi\right] + \sum_{n=1}^{N}\left[\mu_n+\frac{\sigma_n\gamma_n}{\kappa_n}\right]\left[\tau - G(\kappa_n,\tau)\right]\right) \\
&= -\frac{1}{2}\mathrm{Tr}\left[\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\tau\Xi(\tau)\right]\Psi\right] + \sum_{n=1}^{N}\left[\mu_n+\frac{\sigma_n\gamma_n}{\kappa_n}\right]\left[\frac{\mathrm{d}}{\mathrm{d}\tau}\tau - \frac{\mathrm{d}}{\mathrm{d}\tau}G(\kappa_n,\tau)\right] \\
&= -\frac{1}{2}\mathrm{Tr}\left[\Theta(\tau)\Psi\right] + \sum_{n=1}^{N}\left[\mu_n+\frac{\sigma_n\gamma_n}{\kappa_n}\right]\left[1-\exp(-\kappa_n\tau)\right]
\end{aligned} \tag{62}$$

[25] The expressions for $H(\tau)$ and $\tau\Xi_{ij}(\tau)$ arise from substituting the quantities $X = I$, $b_0 = [1,\ldots,1]'$, $a_0 = 0$, $\theta = \mu$, $\kappa = \mathrm{diag}[\kappa_1,\ldots,\kappa_N]$, and $\Sigma = \rho_{ij}\sigma_i\sigma_j$ into the Dai and Singleton (2002) specification. Chen (1995) provides the two-factor result in summation form and Vincente and Tabak (2008) contains the N-factor result in double-summation form, but those expressions are more unweildy. Note also, as discussed in section 3.1, that I have defined the market prices of risk γ_n to be outright positive quantities, so $\frac{\sigma_n\gamma_n}{\kappa_n}$ is added rather than substracted within $H(\tau)$.

where:

$$\begin{aligned}
\Theta_{ij}(\tau) &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left[\tau\Xi(\tau)\right]_{ij} \\
&= \rho_{ij}\sigma_i\sigma_j \cdot \frac{\mathrm{d}}{\mathrm{d}\tau}\left[\tau - G(\kappa_i,\tau) - G(\kappa_j,\tau) + G(\kappa_i+\kappa_j,\tau)\right] \\
&= \rho_{ij}\sigma_i\sigma_j \cdot \left[1 - \exp(-\kappa_i\tau) - \exp(-\kappa_j\tau) + \exp(-\{\kappa_i+\kappa_j\}\tau)\right] \\
&= \rho_{ij}\sigma_i\sigma_j \cdot \left[1 - \exp(-\kappa_i\tau)\right]\left[1 - \exp(-\kappa_j\tau)\right] \\
&= \rho_{ij}\sigma_i\sigma_j \cdot \kappa_i\kappa_j G(\kappa_i,\tau)\, G(\kappa_j,\tau)
\end{aligned} \tag{63}$$

Substituting the expressions $\frac{\mathrm{d}}{\mathrm{d}\tau}G(\kappa_n,\tau)$ and $\frac{\mathrm{d}}{\mathrm{d}\tau}H(\tau)$ into equation 60 gives an expression for the shadow-GATSM forward rate that may be re-arranged to produce the expression in section 3.2, i.e.:

$$\begin{aligned}
\mathrm{f}(t,\tau) &= -\frac{1}{2}\mathrm{Tr}\left[\Theta(\tau)\Psi\right] + \sum_{n=1}^{N}\left[\mu_n + \frac{\sigma_n\gamma_n}{\kappa_n}\right]\left[1 - \exp(-\kappa_n\tau)\right] \\
&\quad + \sum_{n=1}^{N} s_n(t)\exp(-\kappa_n\tau) \\
&= -\frac{1}{2}\mathrm{Tr}\left[\Theta(\tau)\Psi\right] + \sum_{n=1}^{N}\mu_n\left[1 - \exp(-\kappa_n\tau)\right] + s_n(t)\exp(-\kappa_n\tau) \\
&\quad + \sum_{n=1}^{N}\sigma_n\gamma_n \cdot \frac{1}{\kappa_n}\left[1 - \exp(-\kappa_n\tau)\right] \\
&= \sum_{n=1}^{N}\mu_n + \left[s_n(t) - \mu_n\right]\exp(-\kappa_n\tau) + \sum_{n=1}^{N}\sigma_n\gamma_n G(\kappa_n,\tau) \\
&\quad -\frac{1}{2}\mathrm{Tr}\left[\Theta(\tau)\Psi\right]
\end{aligned} \tag{64}$$

B.2 ZLB-GATSM option effect

Chen (1995) p. 360 provides the following closed-form analytic expressions for GATSM option prices:[26]

$$\mathrm{C}(t,\tau,\tau+\delta) = \mathrm{P}(t,\tau+\delta)\,\mathrm{N}\left[d_1(t,\tau,\tau+\delta)\right] - \mathrm{P}(t,\tau)\,\mathrm{N}\left[d_2(t,\tau,\tau+\delta)\right] \tag{65}$$

where:

$$d_1(t,\tau,\tau+\delta) = \frac{1}{\Sigma(\tau,\tau+\delta)}\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right] + \frac{1}{2}\Sigma(\tau,\tau+\delta) \tag{66}$$

$$d_2(t,\tau,\tau+\delta) = d_1(t,\tau,\tau+\delta) - \Sigma(\tau,\tau+\delta) \tag{67}$$

[26] Chen (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.

and:[27]

$$
\begin{aligned}
\left[\Sigma(\tau,\tau+\delta)\right]^2 &= \sum_{m=1}^{N} \sigma_n^2 \left[G(\kappa_n,\delta)\right]^2 G(2\kappa_n,\tau) \\
&\quad + 2\sum_{m=1}^{N}\sum_{n=m+1}^{N} \rho_{mn}\sigma_m\sigma_n G(\kappa_m,\delta)\, G(\kappa_n,\delta)\, G(\kappa_m+\kappa_n,\tau) \qquad (68)
\end{aligned}
$$

I derive the ZLB option effect beginning with its definition in equation 13 and substituting the shadow-GATSM option price expression from equation 65, i.e.:

$$
\begin{aligned}
\mathrm{z}(t,\tau) &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{C}(t,\tau,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\} \\
&= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{P}(t,\tau+\delta)\,\mathrm{N}[d_1(\cdot)] - \mathrm{P}(t,\tau)\,\mathrm{N}[d_2(\cdot)]}{\mathrm{P}(t,\tau)}\right]\right\} \\
&= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\mathrm{N}[d_1(\cdot)] - \mathrm{N}[d_2(\cdot)]\right]\right\} \\
&= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\cdot \mathrm{N}[d_1(\cdot)]\right. \\
&\quad \left. + \frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_1(\cdot)] - \frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_2(\cdot)]\right\} \\
\langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}\cdot \lim_{\delta\to 0}\left\{\mathrm{N}[d_1(t,\tau,\tau+\delta)]\right\} \\
&\quad + \lim_{\delta\to 0}\left\{\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_1(\cdot)] - \frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_2(t,\tau,\tau+\delta)]\right\} \qquad (69)
\end{aligned}
$$

The details for deriving the three limits required for equation 69 are contained in the following subsections, with the following results:

$$
\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\} = -\mathrm{f}(t,\tau) \qquad (70)
$$

$$
\lim_{\delta\to 0}\left\{\mathrm{N}[d_1(t,\tau,\tau+\delta)]\right\} = 1 - \mathrm{N}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right] \qquad (71)
$$

$$
\lim_{\delta\to 0}\left\{\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_1(\cdot)] - \frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_2(\cdot)]\right\} = \omega(\tau)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^2\right) \qquad (72)
$$

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

$$
\mathrm{z}(t,\tau) = -\mathrm{f}(t,\tau)\cdot\left(1 - \mathrm{N}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]\right) + \omega(\tau)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^2\right) \qquad (73)
$$

[27] See 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.

B.2.1 $\lim_{\delta\to 0}\left\{\frac{\mathbf{d}}{\mathbf{d}\delta}\left[\frac{\mathbf{P}(t,\tau+\delta)}{\mathbf{P}(t,\tau)}\right]\right\}$

The expression $\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}$ in equation 69 may be evaluated by first calculating the following derivative :

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

The expression $\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{P}(t,\tau+\delta)$ may be obtained using the result in equation 48 and some re-arrangement, i.e.:

$$\begin{aligned}
\mathrm{f}(t,\tau+\delta) &= -\frac{1}{\mathrm{P}(t,\tau+\delta)}\frac{\mathrm{d}}{\mathrm{d}[\tau+\delta]}\mathrm{P}(t,\tau+\delta) \\
\langle\text{Chain rule}\rangle &= -\frac{1}{\mathrm{P}(t,\tau+\delta)}\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{P}(t,\tau+\delta)\frac{\mathrm{d}\delta}{\mathrm{d}[\tau+\delta]} \\
-\mathrm{f}(t,\tau+\delta)\cdot\mathrm{P}(t,\tau+\delta)\cdot\frac{\mathrm{d}[\tau+\delta]}{\mathrm{d}\delta} &= \frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{P}(t,\tau+\delta) \qquad (74)
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\mathrm{P}(t,\tau)}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{P}(t,\tau+\delta)\right\} &= \frac{1}{\mathrm{P}(t,\tau)}\cdot\lim_{\delta\to 0}\left\{-\mathrm{f}(t,\tau+\delta)\cdot\mathrm{P}(t,\tau+\delta)\cdot\frac{\mathrm{d}[\tau+\delta]}{\mathrm{d}\delta}\right\} \\
&= -\frac{1}{\mathrm{P}(t,\tau)}\cdot\lim_{\delta\to 0}\{\mathrm{f}(t,\tau+\delta)\}\cdot\lim_{\delta\to 0}\{\mathrm{P}(t,\tau+\delta)\cdot 1\} \\
&= -\mathrm{f}(t,\tau) \qquad (75)
\end{aligned}$$

B.2.2 $\lim_{\delta\to 0}\{\mathbf{N}[d_1(t,\tau,\tau+\delta)]\}$

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

$$\mathrm{N}[x] = \frac{1}{2}+\frac{1}{2}\,\mathrm{erf}\left[\frac{1}{\sqrt{2}}x\right] \qquad (76)$$

and the definition of the error function is itself:

$$\mathrm{erf}[y] = \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n[y]^{2n+1}}{n!\,(2n+1)} \qquad (77)$$

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

$$\mathrm{N}[d_1(t,\tau,\tau+\delta)] = \frac{1}{2}+\frac{1}{2}\cdot\frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n\left[\frac{1}{\sqrt{2}}d_1(t,\tau,\tau+\delta)\right]^{2n+1}}{n!\,(2n+1)} \qquad (78)$$

and taking the limit gives:

$$\begin{aligned}
\lim_{\delta\to 0} \mathrm{N}\left[d_1(\cdot)\right] &= \frac{1}{2}+\frac{1}{2}\cdot\lim_{\delta\to 0}\left\{\frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n\left[\frac{1}{\sqrt{2}}d_1(\cdot)\right]^{2n+1}}{n!\,(2n+1)}\right\} \\
\langle\text{Limit rules}\rangle &= \frac{1}{2}+\frac{1}{2}\cdot\frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n\cdot\lim_{\delta\to 0}\left\{\left[\frac{1}{\sqrt{2}}d_1(\cdot)\right]^{2n+1}\right\}}{n!\,(2n+1)} \\
\langle\text{Limit rules}\rangle &= \frac{1}{2}+\frac{1}{2}\cdot\frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n\left[\frac{1}{\sqrt{2}}\cdot\lim_{\delta\to 0}d_1(\cdot)\right]^{2n+1}}{n!\,(2n+1)} \qquad (79)
\end{aligned}$$

where "Limit rules" in the final line includes the power limit rule $\lim_{x\to a}\left\{\left[f(x)\right]^n\right\} = \left\{\left[\lim_{x\to a} f(x)\right]\right\}^n$.

The expression for $\lim_{\delta\to 0} d_1(t,\tau,\tau+\delta)$ is derived directly from the expression in equation 66, i.e.:

$$\begin{aligned}
\lim_{\delta\to 0} d_1(t,\tau,\tau+\delta) &= \lim_{\delta\to 0}\left\{\frac{1}{\Sigma(\tau,\tau+\delta)}\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]+\frac{1}{2}\Sigma(\tau,\tau+\delta)\right\} \\
\langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{\frac{1}{\Sigma(\tau,\tau+\delta)}\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}+\frac{1}{2}\lim_{\delta\to 0}\Sigma(\tau,\tau+\delta) \\
\langle\text{L'Hopital's rule}\rangle &= \frac{\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}}{\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\tau,\tau+\delta)\right\}}+\frac{1}{2}\lim_{\delta\to 0}\Sigma(\tau,\tau+\delta) \qquad (80)
\end{aligned}$$

where L'Hopital's rule has been applied to the first expression on the second line of equation 80 because $\lim_{\delta\to 0}\Sigma(\tau,\tau+\delta)$ is readily calculated as zero, immediately below, meaning the limit would otherwise be undefined (i.e. $\log\left[\mathrm{P}(t,\tau)/\mathrm{P}(t,\tau)\right] = \log[1] = 0$ and $\Sigma(\tau,\tau) = 0$).

Specifically regarding $\lim_{\delta\to 0}\Sigma(\tau,\tau+\delta)$, referring to equation 68 and noting that for any x:

$$\begin{aligned}
\lim_{\delta\to 0} G(x,\delta) &= \lim_{\delta\to 0}\left\{\frac{1-\exp(-x\delta)}{x}\right\} \\
&= \frac{1-1}{x} \\
&= 0 \qquad (81)
\end{aligned}$$

directly gives the result:

$$\lim_{\delta\to 0}\Sigma(\tau,\tau+\delta) = 0 \qquad (82)$$

The expression $\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\}$ required for equation 80 is:

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right\} &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left(\log\left[\mathrm{P}(t,\tau+\delta)\right]-\log\left[\mathrm{P}(t,\tau)\right]\right)\right\}\\
&= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\log\left[\mathrm{P}(t,\tau+\delta)\right]-\frac{\mathrm{d}}{\mathrm{d}\delta}\log\left[\mathrm{P}(t,\tau)\right]\right\}\\
\langle\text{Chain rule}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\left[\tau+\delta\right]}\log\left[\mathrm{P}(t,\tau+\delta)\right]\cdot\frac{\mathrm{d}\left[\tau+\delta\right]}{\mathrm{d}\delta}\right\}\\
&= \lim_{\delta\to 0}\left\{-\mathrm{f}(t,\tau+\delta)\cdot 1\right\}\\
&= -\mathrm{f}(t,\tau)
\end{aligned}\tag{83}$$

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

$$\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\tau,\tau+\delta)\right\}=\omega(\tau)\tag{84}$$

Substituting these results into equation 80 gives:

$$\lim_{\delta\to 0}d_1(t,\tau,\tau+\delta)=-\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\tag{85}$$

and the final result for $\lim_{\delta\to 0}\mathrm{N}[d_1(t,\tau,\tau+\delta)]$ is obtained substituting that result into equation 79, i.e.:

$$\begin{aligned}
\lim_{\delta\to 0}\mathrm{N}\left[d_1(t,\tau,\tau+\delta)\right] &= \frac{1}{2}+\frac{1}{2}\cdot\frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n\left[\frac{1}{\sqrt{2}}\cdot\left(-\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right)\right]^{2n+1}}{n!\,(2n+1)}\\
&= \frac{1}{2}+\frac{1}{2}\cdot\mathrm{erf}\left[-\frac{1}{\sqrt{2}}\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]\\
&= \mathrm{N}\left[-\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]\\
&= 1-\mathrm{N}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]
\end{aligned}\tag{86}$$

B.2.3 $\lim_{\delta\to 0}\left\{\frac{\mathbf{P}(t,\tau+\delta)}{\mathbf{P}(t,\tau)}\frac{\mathbf{d}}{\mathbf{d}\delta}\mathbf{N}\left[d_1(t,\tau,\tau+\delta)\right]-\frac{\mathbf{d}}{\mathbf{d}\delta}\mathbf{N}\left[d_2(t,\tau,\tau+\delta)\right]\right\}$

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

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}\left[d_1(t,\tau,\tau+\delta)\right] &= \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{N}[x]\,\frac{\mathrm{d}}{\mathrm{d}\delta}d_1(t,\tau,\tau+\delta)\\
\langle\text{Chain rule}\rangle &: \left\langle x=d_1(t,\tau,\tau+\delta)\,;\ \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{N}[x]=\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}x^2\right]\right\rangle\\
&= \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left[d_1(t,\tau,\tau+\delta)\right]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}d_1(t,\tau,\tau+\delta)
\end{aligned}\tag{87}$$

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

$$\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}\left[d_1(t,\tau,\tau+\delta)\right] = \frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left[d_1(\cdot)\right]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}d_1(\cdot) \tag{88}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}\left[d_2(t,\tau,\tau+\delta)\right] = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left[d_2(t,\tau,\tau+\delta)\right]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}d_2(t,\tau,\tau+\delta) \tag{89}$$

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

$$\begin{aligned}
\left[d_2(t,\tau,\tau+\delta)\right]^2 &= \left[d_1(t,\tau,\tau+\delta) - \Sigma(\tau,\tau+\delta)\right]^2 \\
&= \left[d_1(\cdot)\right]^2 - 2d_1(\cdot)\Sigma(\tau,\tau+\delta) + \left[\Sigma(\tau,\tau+\delta)\right]^2 \\
&= \left[d_1(\cdot)\right]^2 - 2\left(\frac{1}{\Sigma(\cdot)}\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right] + \frac{1}{2}\Sigma(\cdot)\right)\Sigma(\cdot) - \left[\Sigma(\cdot)\right]^2 \\
&= \left[d_1(\cdot)\right]^2 - 2\cdot\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right] + \left[\Sigma(\cdot)\right]^2 - \left[\Sigma(\cdot)\right]^2 \\
&= \left[d_1(\cdot)\right]^2 - 2\cdot\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]
\end{aligned} \tag{90}$$

and therefore:

$$\begin{aligned}
\exp\left[-\frac{1}{2}\left[d_2(t,\tau,\tau+\delta)\right]^2\right] &= \exp\left[-\frac{1}{2}\left(\left[d_1(\cdot)\right]^2 - 2\cdot\log\left[\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right]\right)\right] \\
&= \exp\left[-\frac{1}{2}\left[d_1(\cdot)\right]^2\right]\cdot\exp\left[\log\left(\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right)\right] \\
&= \frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\exp\left[-\frac{1}{2}\left[d_1(t,\tau,\tau+\delta)\right]^2\right]
\end{aligned} \tag{91}$$

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

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\delta}d_2(t,\tau,\tau+\delta) &= \frac{\mathrm{d}}{\mathrm{d}\delta}\left[d_1(t,\tau,\tau+\delta) - \Sigma(\tau,\tau+\delta)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}\delta}d_1(t,\tau,\tau+\delta) - \frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\tau,\tau+\delta)
\end{aligned} \tag{92}$$

Substituting the results from equations 91 and 92 into equation 89 gives the following expression:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}\left[d_2(\cdot)\right] &= \frac{1}{\sqrt{2\pi}}\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\exp\left[-\frac{1}{2}\left[d_1(\cdot)\right]^2\right]\left[\frac{\mathrm{d}}{\mathrm{d}\delta}d_1(\cdot) - \frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\tau,\tau+\delta)\right] \\
&= \frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left[d_1(\cdot)\right]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}d_1(\cdot) \\
&\quad -\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left[d_1(\cdot)\right]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\tau,\tau+\delta)
\end{aligned} \tag{93}$$

When calculating $U(t,\tau,\tau+\delta) = \frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_1(t,\tau,\tau+\delta)] - \frac{\mathrm{d}}{\mathrm{d}\delta}\mathrm{N}[d_2(t,\tau,\tau+\delta)]$, note that the expression $\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}[d_1(\cdot)]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}d_1(\cdot)$ in equation 88 cancels identically with the first component from equation 93. Therefore:

$$U(t,\tau,\tau+\delta) = \frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}[d_1(\cdot)]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\cdot) \tag{94}$$

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

$$\begin{aligned}
\lim_{\delta\to 0} U(t,\tau,\tau+\delta) &= \lim_{\delta\to 0}\left\{\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}[d_1(\cdot)]^2\right]\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\tau,\tau+\delta)\right\} \\
\langle\text{Limit rules}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right\}\cdot\lim_{\delta\to 0}\left\{\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}[d_1(\cdot)]^2\right]\right\} \\
&\quad\times\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma(\tau,\tau+\delta)\right\}
\end{aligned} \tag{95}$$

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

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{\mathrm{P}(t,\tau+\delta)}{\mathrm{P}(t,\tau)}\right\} &= \frac{\mathrm{P}(t,\tau)}{\mathrm{P}(t,\tau)} \\
&= 1
\end{aligned} \tag{96}$$

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

$$\exp[x] = \sum_{n=0}^{\infty}\frac{x^n}{n!} \tag{97}$$

Substituting $x = -\frac{1}{2}[d_1(t,\tau,\tau+\delta)]^2$ into equation 97 and taking the limit gives:

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}[d_1(\cdot)]^2\right]\right\} &= \frac{1}{\sqrt{2\pi}}\lim_{\delta\to 0}\left\{\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}[d_1(\cdot)]^2\right)^n}{n!}\right\} \\
&= \frac{1}{\sqrt{2\pi}}\lim_{\delta\to 0}\left\{\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^n[d_1(\cdot)]^{2n}}{n!}\right\} \\
\langle\text{Limit rules}\rangle &= \frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^n\lim_{\delta\to 0}\left\{[d_1(\cdot)]^{2n}\right\}}{n!} \\
\langle\text{Limit rules}\rangle &= \frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^n\left[\lim_{\delta\to 0}d_1(t,\tau,\tau+\delta)\right]^{2n}}{n!}
\end{aligned} \tag{98}$$

The result $\lim_{\delta\to 0}\left[d_1\left(t,\tau,\tau+\delta\right)\right] = -\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}$ is available from equation 85, and so

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left[d_1\left(t,\tau,\tau+\delta\right)\right]^2\right]\right\} &= \frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^n\left[-\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^{2n}}{n!} \\
&= \frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^n\left(\left[-\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^{2}\right)^n}{n!} \\
&= \frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\left[\frac{\mathrm{f}(t,\tau)}{\omega(\tau)}\right]^{2}\right)^n}{n!} \\
&= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}\left(t,\tau\right)}{\omega\left(\tau\right)}\right]^2\right) \qquad (99)
\end{aligned}$$

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

$$\begin{aligned}
\lim_{\delta\to 0} U\left(t,\tau,\tau+\delta\right) &= 1\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}\left(t,\tau\right)}{\omega\left(\tau\right)}\right]^2\right)\cdot\omega\left(\tau\right) \\
&= \omega\left(\tau\right)\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{\mathrm{f}\left(t,\tau\right)}{\omega\left(\tau\right)}\right]^2\right) \qquad (100)
\end{aligned}$$

B.3 Annualized instantaneous volatility

I define annualized instantaneous volatility $\omega\left(\tau\right)$ as the annualized limit of the option volatility expression from equation 68 and denote it $\omega\left(\tau\right)$, i.e.:

$$\begin{aligned}
\omega\left(\tau\right) &= \lim_{\delta\to 0}\left\{\frac{1}{\delta}\Sigma\left(\tau,\tau+\delta\right)\right\} \\
\langle\text{L'Hopital's rule}\rangle &= \frac{\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma\left(\tau,\tau+\delta\right)\right\}}{\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\delta\right\}} \\
&= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma\left(\tau,\tau+\delta\right)\right\} \qquad (101)
\end{aligned}$$

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

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

$$
\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma\left(\tau,\tau+\delta\right)\right\} &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\sqrt{\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2}\right\} \\
\langle\text{Chain rule}\rangle &= \lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{x}\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2\right\} \\
&\quad \left\langle x=\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2;\ \frac{\mathrm{d}}{\mathrm{d}x}\sqrt{x}=\frac{1}{2}x^{-1/2}=\frac{1}{2\sqrt{x}}\right\rangle \\
&= \lim_{\delta\to 0}\left\{\frac{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2}{2\Sigma\left(\tau,\tau+\delta\right)}\right\} \\
\langle\text{L'Hopital's rule}\rangle &= \frac{\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2\right]\right\}}{2\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma\left(\tau,\tau+\delta\right)\right\}} \\
&= \frac{\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2\right\}}{2\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma\left(\tau,\tau+\delta\right)\right\}}
\end{aligned}
\tag{102}
$$

Note that L'Hopital's rule has been used in the second-last step of equation 102 because in the middle line both $\frac{\mathrm{d}}{\mathrm{d}\delta}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2=2\Sigma\left(\tau,\tau+\delta\right)\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma\left(\tau,\tau+\delta\right)$ and $2\Sigma\left(\tau,\tau+\delta\right)$ would equal zero when evaluated at $\delta=0$ (see equation 82), which would leave the limit undefined.

Re-arranging equation 102 gives:

$$\left[\lim_{\delta\to 0}\left\{\frac{\mathrm{d}}{\mathrm{d}\delta}\Sigma\left(\tau,\tau+\delta\right)\right\}\right]^2=\frac{1}{2}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2\right\} \tag{103}$$

and therefore:

$$\omega\left(\tau\right)=\sqrt{\frac{1}{2}\cdot\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2\right\}} \tag{104}$$

Referring to equation 68, $\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2\right\}$ in equation 104 may be evaluated by calculating the limits of the double derivatives of the functions $\left[G\left(\kappa_n,\delta\right)\right]^2$ and $G\left(\kappa_m,\delta\right)G\left(\kappa_n,\delta\right)$ and then substituting those results into equation 68. That is:

$$
\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[G\left(\kappa_n,\delta\right)\right]^2 &= \frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{d}}{\mathrm{d}\delta}\left[G\left(\kappa_n,\delta\right)\right]^2\right] \\
\langle\text{Chain rule}\rangle &= \frac{\mathrm{d}}{\mathrm{d}\delta}\left[2G\left(\kappa_n,\delta\right)\right]\exp\left(-\kappa_n\delta\right) \\
&= 2\frac{\mathrm{d}}{\mathrm{d}\delta}G\left(\kappa_n,\delta\right)\cdot\exp\left(-\kappa_n\delta\right)+2G\left(\kappa_n,\delta\right)\frac{\mathrm{d}}{\mathrm{d}\delta}\exp\left(-\kappa_n\delta\right) \\
&= 2\exp\left(-\kappa_n\delta\right)\cdot\exp\left(-\kappa_n\delta\right)+2G\left(\kappa_n,\delta\right)\cdot-\kappa_n\exp\left(-\kappa_n\delta\right) \\
&= 2\exp\left(-2\kappa_n\delta\right)-2G\left(\kappa_n,\delta\right)\cdot\kappa_n\exp\left(-\kappa_n\delta\right)
\end{aligned}
\tag{105}
$$

which has a limit:

$$\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[G\left(\kappa_n,\delta\right)\right]^2\right\}=2-0=2 \tag{106}$$

and:

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[G\left(\kappa_m,\delta\right)G\left(\kappa_n,\delta\right)\right] &= \frac{\mathrm{d}}{\mathrm{d}\delta}\left[\frac{\mathrm{d}}{\mathrm{d}\delta}\left[G\left(\kappa_m,\delta\right)G\left(\kappa_n,\delta\right)\right]\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}\delta}\left[\exp\left(-\kappa_m\delta\right)G\left(\kappa_n,\delta\right)+G\left(\kappa_m,\delta\right)\exp\left(-\kappa_n\delta\right)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}\delta}\exp\left(-\kappa_m\delta\right)\cdot G\left(\kappa_n,\delta\right)+\exp\left(-\kappa_m\delta\right)\frac{\mathrm{d}}{\mathrm{d}\delta}G\left(\kappa_n,\delta\right) \\
&\quad +\frac{\mathrm{d}}{\mathrm{d}\delta}G\left(\kappa_m,\delta\right)\cdot\exp\left(-\kappa_n\delta\right) \\
&\quad +G\left(\kappa_m,\delta\right)\frac{\mathrm{d}}{\mathrm{d}\delta}\exp\left(-\kappa_n\delta\right) \\
&= -\kappa_m\exp\left(-\kappa_m\delta\right)G\left(\kappa_n,\delta\right)+\exp\left(-\kappa_m\delta\right)\exp\left(-\kappa_n\delta\right) \\
&\quad +\exp\left(-\kappa_m\delta\right)\exp\left(-\kappa_n\delta\right) \\
&\quad +G\left(\kappa_m,\delta\right)\cdot-\kappa_n\exp\left(-\kappa_n\delta\right) \\
&= -\kappa_m\exp\left(-\kappa_m\delta\right)G\left(\kappa_n,\delta\right)+2\exp\left(-\left\{\kappa_m+\kappa_n\right\}\delta\right) \\
&\quad -G\left(\kappa_m,\delta\right)\kappa_n\exp\left(-\kappa_n\delta\right) \qquad (107)
\end{aligned}$$

which has a limit:

$$\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[G\left(\kappa_m,\delta\right)G\left(\kappa_n,\delta\right)\right]\right\}=0+2+0=2 \qquad (108)$$

Substituting the two limit results into equation 68 $\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[\Sigma\left(\tau,\tau+\delta\right)\right]^2\right\}$ gives:

$$\begin{aligned}
\lim_{\delta\to 0}\left\{\frac{\mathrm{d}^2}{\mathrm{d}\delta^2}\left[\Sigma\left(\cdot\right)\right]^2\right\} &= \sum_{n=1}^{N}2\sigma_n^2\cdot G\left(2\kappa_n,\tau\right) \\
&\quad +2\sum_{m=1}^{N}\sum_{n=m+1}^{N}2\rho_{mn}\sigma_m\sigma_n\cdot G\left(\kappa_m+\kappa_n,2\tau\right) \qquad (109)
\end{aligned}$$

and substituting the preceding result into equation 104 gives:

$$\omega\left(\tau\right)=\sqrt{\sum_{n=1}^{N}\sigma_n^2\cdot G\left(2\kappa_n,\tau\right)+2\sum_{m=1}^{N}\sum_{n=m+1}^{N}\rho_{mn}\sigma_m\sigma_n\cdot G\left(\kappa_m+\kappa_n,\tau\right)} \qquad (110)$$

Note that I have retained a form for $\omega\left(\tau\right)$ that is analogous to Chen (1995); $\omega\left(\tau\right)$ could be expressed more compactly as:

$$\omega\left(\tau\right)=\sqrt{\sum_{m=1}^{N}\sum_{n=1}^{N}\rho_{mn}\sigma_m\sigma_n\cdot G\left(\kappa_m+\kappa_n,\tau\right)} \qquad (111)$$

with $\rho_{nn}=1$. Indeed, the double-summation itself could be expressed more compactly as the sum of the matrix elements $\rho_{ij}\sigma_i\sigma_j\cdot G\left(\kappa_i+\kappa_j,\tau\right)$.

C Generic GATSM to GATSM(2)

This appendix details how the shadow-GATSM(2) outlined in section 4.1 may be obtained using the generic shadow-GATSM expression outlined in section 3.2.

I first set $N = 2$ in equation 27, and then set $\mu_1 = \mu_2 = 0$. The resulting interim expression is:

$$\mathrm{f}(t,\tau) = -\frac{1}{2}\mathrm{Tr}\left[\Theta(\tau)\Psi\right] + \sum_{n=1}^{2} s_n(t)\cdot\exp(-\kappa_n\tau) + \sum_{n=1}^{2}\sigma_n\gamma_n\cdot G(\kappa_n,\tau) \tag{112}$$

Regarding the volatility effect component $-\frac{1}{2}\mathrm{Tr}[\Theta(\tau)\Psi]$, $\Theta_{ij}(\tau)$ in explicit matrix form is:

$$\Theta(\tau) = \begin{bmatrix} \sigma_1^2\cdot\kappa_1^2\left[G(\kappa_1,\tau)\right]^2 & \rho\sigma_1\sigma_2\cdot\kappa_1\kappa_2 G(\kappa_1,\tau)\,G(\kappa_2,\tau) \\ \rho\sigma_1\sigma_2\cdot\kappa_1\kappa_2 G(\kappa_1,\tau)\,G(\kappa_2,\tau) & \sigma_2^2\cdot\kappa_2^2\left[G(\kappa_2,\tau)\right]^2 \end{bmatrix} \tag{113}$$

and Ψ_{ij} is:

$$\Psi = \begin{bmatrix} \frac{1}{\kappa_1^2} & \frac{1}{\kappa_1\kappa_2} \\ \frac{1}{\kappa_1\kappa_2} & \frac{1}{\kappa_2^2} \end{bmatrix} \tag{114}$$

Therefore, the trace evaluation in equation 112 is:

$$\begin{aligned}\mathrm{Tr}\left[\Theta(\tau)\Psi\right] &= \mathrm{Tr}\left\{\mathrm{diag}\begin{bmatrix} \sigma_1^2\cdot\left[G(\kappa_1,\tau)\right]^2 + \rho\sigma_1\sigma_2\cdot G(\kappa_1,\tau)\,G(\kappa_2,\tau) \\ \sigma_2^2\cdot\left[G(\kappa_2,\tau)\right]^2 + \rho\sigma_1\sigma_2\cdot G(\kappa_1,\tau)\,G(\kappa_2,\tau) \end{bmatrix}\right\} \\ &= \sigma_1^2\cdot\left[G(\kappa_1,\tau)\right]^2 + \sigma_2^2\cdot\left[G(\kappa_2,\tau)\right]^2 \\ &\qquad + 2\rho\sigma_1\sigma_2\cdot G(\kappa_1,\tau)\,G(\kappa_2,\tau)\end{aligned} \tag{115}$$

where the expression $\mathrm{Tr}\{\mathrm{diag}[\cdot]\}$ recognizes that the off-diagonal elements of $\Theta(\tau)\Psi$ are irrelevant. Combining the explicit volatility effect with the remainder of equation 112 gives the following updated expression for the forward rate:

$$\begin{aligned}\mathrm{f}(t,\tau) &= s_1(t)\cdot\exp(-\kappa_1\tau) + s_2(t)\cdot\exp(-\kappa_2\tau) \\ &\quad + \sigma_1\gamma_1\cdot G(\kappa_1,\tau) + \sigma_2\gamma_2\cdot G(\kappa_2,\tau) \\ &\quad - \sigma_1^2\cdot\frac{1}{2}\left[G(\kappa_1,\tau)\right]^2 - \sigma_2^2\cdot\frac{1}{2}\left[G(\kappa_2,\tau)\right]^2 \\ &\qquad - \rho\sigma_1\sigma_2\cdot G(\kappa_1,\tau)\,G(\kappa_2,\tau)\end{aligned} \tag{116}$$

The expression for annualized instantaneous volatility from equation 31 with two factors is:

$$\omega(\tau) = \sqrt{\sigma_1^2\cdot G(2\kappa_1,\tau) + \sigma_2^2\cdot G(2\kappa_2,\tau) + 2\rho\sigma_1\sigma_2\cdot G(\kappa_1,\tau)\,G(\kappa_2,\tau)} \tag{117}$$

Both $\mathrm{f}(t,\tau)$ and $\omega(\tau)$ can be made more parsimonious by taking the limit as $\kappa_1 \to 0$.

Specifically, $\lim_{\kappa_1 \to 0} \exp(-\kappa_1 \tau) = 1$, and:

$$\begin{aligned} \lim_{\kappa_1 \to 0} G(\kappa_1, \tau) &= \lim_{\kappa_1 \to 0} \left\{ \frac{1}{\kappa_1} [1 - \exp(-\kappa_1 \tau)] \right\} \\ \langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\kappa_1 \to 0} \left\{ \frac{\mathrm{d}}{\mathrm{d}\kappa_1} [1 - \exp(-\kappa_1 \tau)] \right\}}{\lim_{\kappa_1 \to 0} \left\{ \frac{\mathrm{d}}{\mathrm{d}\kappa_1} \kappa_1 \right\}} \\ &= \frac{\lim_{\kappa_1 \to 0} \{\tau \exp(-\kappa_1 \tau)\}}{\lim_{\kappa_1 \to 0} \{1\}} \\ &= \tau \end{aligned}$$

while $\lim_{\kappa_1 \to 0} \left\{ \frac{1}{2} [G(\kappa_1, \tau)]^2 \right\} = \frac{1}{2}\tau^2$ and $\lim_{\kappa_1 \to 0} [G(2\kappa_1, \tau)]^2 = \tau^2$, which is evident from the following limit evaluation:

$$\begin{aligned} \lim_{x \to 0} \left\{ [G(x, \tau)]^2 \right\} &= \left[\lim_{x \to 0} G(x, \tau) \right]^2 \\ &= \tau^2 \end{aligned}$$

Substituting the limit results into equations 116 and 117, and setting $\kappa_2 = \lambda$ for notational convenience gives the ZLB-GATSM(2) results for $\mathrm{f}(t, \tau)$ and $\omega(\tau)$ provided in section 4.1.

D Very-long-horizon ZLB term structure example

The ZLB-GATSM(2) example in this appendix shows how the ZLB-GATSM framework accommodates negative shadow forward rates and interest rates that can arise for very long maturities in shadow-GATSMs.

As for figure 3 in section 4.2, I set the ZLB-GATSM(2) state variables to $s_1(t) = 0.05$ and $s_2(t) = 0$, which results in a value for the current shadow short rate of $\mathrm{r}(t) = s_1(t) + s_2(t) = 0.05$, or five percent. With $s_2(t) = 0$, the expected path of the short rate $\mathbb{E}_t[\mathrm{r}(t + \tau)]$ is also five percent for all horizons, although the long-horizon results being investigated here 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 very long times to maturity (beyond around 30 years in this example), it is a mathematical inevitability that the shadow volatility effect term will become large enough to result in negative shadow forward rates. The negative values result from the increasing magnitude of the time-to-maturity quadratic term $-\sigma_1^2 \cdot \frac{1}{2}\tau^2$ in the volatility effect component $-\sigma_1^2 \cdot \frac{1}{2}\tau^2 - \sigma_2^2 \cdot \frac{1}{2} [G(\lambda, \tau)]^2$ for the shadow-GATSM(2). Being a cumulative average, negative shadow interest rates also eventually result (beyond around 50 years in this example).

The ZLB-GATSM framework resolves the issue of negative very-long-horizon forward rates and interest rates, technically at least, as illustrated in the top-right subplot of figure 3. That is, the option effects take on increasingly positive values that offset the increasingly negative long-horizon forward rates and interest rates. The very long horizon in figure 3 (i.e. 80 years) shows that forward rates eventually asymptote to zero. ZLB interest rates also asymptote to zero, but do so much more slowly because the standard term structure relationship $\underline{\mathrm{r}}(t,\tau) = \frac{1}{\tau}\int_0^\tau \underline{\mathrm{f}}(t,\upsilon)\mathrm{d}\upsilon$ effectively becomes a reciprocal function of time to maturity once $\underline{\mathrm{f}}(t,\upsilon) \simeq 0$ stops contributing to $\int_0^\tau \underline{\mathrm{f}}(t,\upsilon)\mathrm{d}\upsilon$. For example, $\underline{\mathrm{r}}(t,100) = 3.58$ percent and $\underline{\mathrm{r}}(t,200) = \underline{\mathrm{r}}\,(t,100)\,/2 = 1.79$ percent, etc.

Of course, whether zero represents plausible values for very-long-horizon forward rates is an open question, but the ZLB-GATSM framework does at least deliver non-negative forward rates and interest rates all horizons/times to maturity.

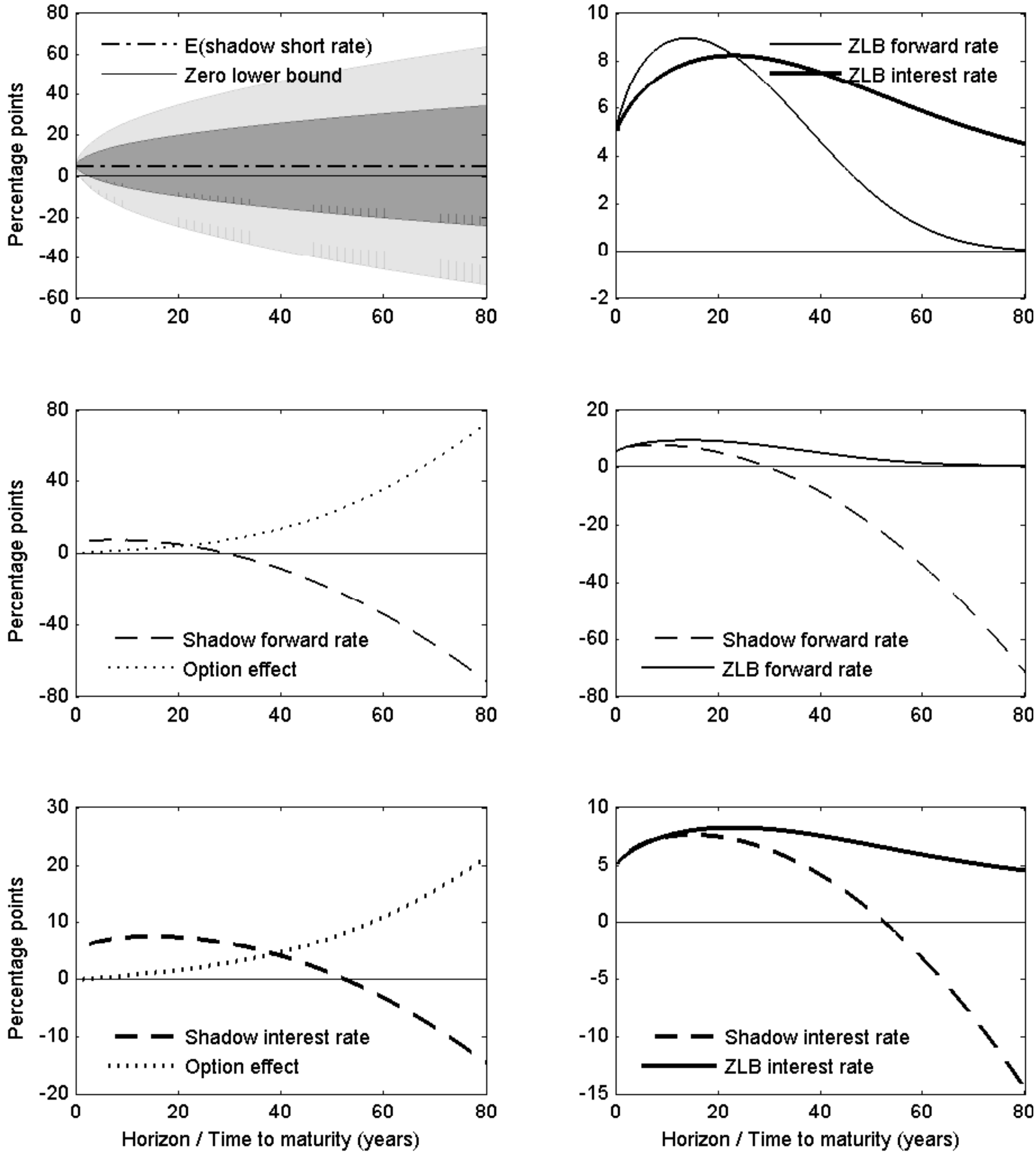

Figure 9: Perspectives on the ZLB-GATSM(2) term structure with $s_1(t) = 0.05$ and $s_1(t) = 0$, so $\mathrm{r}(t)$ and $\mathbb{E}_t[\mathrm{r}(t+\tau)] = 5$ percent.