



DP2013/02

**A tractable framework for zero lower bound
Gaussian term structure models**

Leo Krippner

August 2013

JEL classification: E43, G12, G13

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

Discussion Paper Series

ISSN 1177-7567

DP2013/02

**A tractable framework for zero lower bound Gaussian
term structure models***

Leo Krippner[†]

Abstract

When nominal interest rates are near their zero lower bound (ZLB), as in many developed economies at the time of writing, it is theoretically untenable to apply the popular class of Gaussian affine term structure models (GATSMs) given their inherent material probabilities of negative interest rates. Hence, I propose a new and tractable modification for GATSMs that enforces the ZLB, and which approximates the fully arbitrage-free but much less tractable framework proposed in Black (1995). I apply my framework to United States yield curve data, with robust estimation via the iterated extended Kalman filter, and first show that the two-factor results are very similar to those from a comparable Black model. I then estimate two- and three-factor models with longer-maturity data sets to illustrate that my ZLB framework can readily be applied in circumstances that would be computationally burdensome or infeasible within the Black framework.

* 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. This paper supersedes the RBNZ Discussion Paper DP2012/02, "Modifying Gaussian term structure models when interest rates are near the zero lower bound". For their many helpful comments, the author thanks colleagues at the Reserve Bank, Iris Claus, Toby Daghish, Francis Diebold, Pedro Gomis, Maurice Obstfeld, Glenn Rudebusch, Peter Phillips, Yoichi Ueno, and participants at the following presentations: the 2011 Reserve Bank of New Zealand conference, the 2012 New Zealand Econometrics Study Group meeting, the 2013 mid-Western Finance Association meeting, the 2013 New Zealand Association of Economists conference, and seminars at the Bundesbank, the Bank of Japan, the Federal Reserve Board, and the Federal Reserve Banks of St. Louis, San Francisco, Chicago, and New York. I especially thank James Bullard for his encouragement on the topic, and Scott Richard for his observations that led to section 4.4.

[†] 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 ©Reserve Bank of New Zealand

1 Introduction

In this article I propose a new and tractable framework for imposing the zero lower bound (hereafter ZLB) for nominal interest rates on Gaussian affine term structure models (hereafter GATSMs).

Ruling out negative interest rates in term structure models is desirable on theoretical and practical grounds, as I expand on subsequently below, and I use GATSMs as the basis for the framework because they are very flexible and tractable. Specifically, GATSMs with any number of factors and any factor inter-relationships retain 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 apply and estimate relative to other term structure models. As such, Hamilton and Wu (2012) introduces GATSMs as “the basic workhorse in macroeconomics and finance” and notes fourteen recent examples of their application. Rudebusch (2010) surveys GATSM applications in macro-finance, and Diebold and Rudebusch (2013) provides an overview of many forecasting, finance, and macro-finance applications.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. If negative interest rates did occur, one could realize an arbitrage profit by borrowing funds (therefore receiving the absolute interest rate) and holding them as 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 theoretical inconsistency, GATSMs are often applied with the assumption that the inherent probabilities of negative interest rates in GATSMs are sufficiently small to make such models immaterially different to a “real world” model subject to the ZLB.² While that assumption may have been tenable historically, when interest rate levels relative to their typical volatilities remained well above zero, it is obviously untenable in many developed economies at the time of writing. For example, near-zero policy interest rates have been maintained in Japan since the 1990s, in the United States and the United Kingdom since late-2008/early-2009, and in the euro area and non-euro European economies over recent years. Moreover, the practical constraint of the ZLB is evidenced by many of those countries having adopted unconventional monetary policy easing (e.g. large-scale asset purchases and/or explicit guidance on future policy rates) to circumvent the constraint on easing monetary policy via the conventional method of lowering interest rates. The ZLB constraint is also apparent in the short- and mid-maturity interest rates of many of the countries already mentioned, being at levels well within their typical historical volatilities.

¹Diebold and Rudebusch (2013) summarizes the development and application of arbitrage-free versions of the Nelson and Siegel (1987) class, which are an empirically reliable sub-class of GATSMs. Joslin, Singleton, and Zhu (2011) provides an empirically reliable method for estimating GATSMs in general, and Hamilton and Wu (2012) provides another approach.

²See, for example, Piazzesi (2010) p. 716 or Filipović (2009) p. 86. However, the assumption is typically left implicit.

In near-zero interest rate environments, the material probability of negative interest rates in GATSMs 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. Such mis-specifications will affect even routine GATSM applications, such as monitoring the level and shape of the estimated term structure to provide a gauge of the stance of monetary policy. The implications of such mis-specifications may be further compounded for any macro-finance relationships established between GATSMs and macroeconomic data (e.g. measures of inflation and real output growth), because macroeconomic data are not constrained to be non-negative.

Black (1995, hereafter Black) provides a straightforward and theoretically appealing framework that can be used in conjunction with GATSMs to eliminate negative interest rates.³ The Black framework is based on the observation that physical currency effectively provides an option against negative interest rates at each point in time. Specifically, the Black short rate may be defined as $\underline{r}(t) = \max\{r(t), 0\}$, where $r(t)$ 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 $r(t)$ within the Black framework) result in models with limited tractability. For example, even the simplest Black-GATSM, the one-factor Black-Vasicek (1977) model investigated in Gorovoi and Linetsky (2004), does not result in a closed-form analytic solution. Specifically, it requires the numerical evaluation of relatively complex Weber-Hermite parabolic cylinder functions and numerical integration. In addition, the Gorovoi and Linetsky (2004) approach does not appear to generalize to multiple factors.⁴ To implement two-factor Black-GATSMs, Bomfim (2003), Ueno, Baba, and Sakurai (2006), Ichiue and Ueno (2007), Kim and Singleton (2012), and Ichiue and Ueno (2013) variously apply the purely numerical methods of finite-difference grids, interest rate lattices, and Monte Carlo simulations (to US, Japanese, and UK data). Bauer and Rudebusch (2013) and Richard (2013) are examples of three-factor Black-GATSM applications, but they also illustrate the practical challenge of applying numerical methods for higher dimensional models.⁵ In general,

³Of 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. 226-33 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; (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 (1985)/square-root models, and even the special case of independent factors requires the relatively complex evaluation of the product of noncentral chi-square distributions (e.g. see Chen and Scott (1992) for the two-factor case and the comment in Piazzesi (2010) p. 727 in general); and (3) Such models may imply unrealistic interest rate dynamics near the ZLB. For example, Cox, Ingersoll, and Ross (1985)/square-root models have the inherent property that interest rate innovations approach zero as interest rates themselves approach zero, which is counter to empirical evidence.

⁴See Kim and Singleton (2012) p. 11. Ichiue and Ueno (2006) and Ueno, Baba, and Sakurai (2006) apply the Gorovoi and Linetsky (2004) framework to the Japanese term structure.

⁵Bauer and Rudebusch (2013), which uses Monte Carlo simulation and data out to 10 years time

the number of evaluations required for numerical methods increases as the power of the number of factors and precision declines for longer times to maturity.⁶ Therefore, while Black-GATSMs have arbitrary flexibility in principle, in practice their rapid decline in tractability for more than two factors and longer times to maturity would make implementation computationally burdensome or practically infeasible. Therefore, the Black framework cannot in practice compete with the full flexibility that is currently taken for granted when applying multifactor GATSMs to data sets of any maturity span.

The alternative and tractable ZLB framework that I propose in this article is based on the Black foundation of the optionality provided by physical currency. However, by imposing a slight approximation to Black, as I will define precisely in sections 2 and 3 in light of the appropriate notation and background, my ZLB framework in the Gaussian context results in simple closed-form expressions for the physical currency option effect as a function of time to maturity. Adding that expression to the shadow-GATSM forward rate curve results in a ZLB forward rate curve with a simple closed-form analytic expression. The result is general and so applies to any shadow-GATSM specification. Hence, irrespective of the GATSM used to represent the shadow short term structure, ZLB forward rates will always be a function of just scalar exponential functions and the standard cumulative normal function. ZLB-GATSM interest rates and ZLB-GATSM bond prices are obtained using standard term structure relationships; respectively, 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 the nature of the ZLB-GATSM forward rate expression guarantees that it will always be univariate (with respect to time to maturity) and therefore elementary. In turn, the tractability of my ZLB-GATSM framework combined with multivariate-normal transition densities for the state variables facilitates the estimation of ZLB models with yield curve data that is materially constrained by the ZLB. Note that since my framework was introduced in an earlier version of the present article, it has already been applied in Krippner (2013) and Christensen and Rudebusch (2013), and used in Bullard (2012, 2013). Wu and Xia (2013) has subsequently developed and applied the equivalent framework based on a discrete-time shadow-GATSM.

Notwithstanding the relative benefit of tractability offered by my ZLB-GATSM framework, the cost is that it is not a theoretically self-consistent (i.e. fully arbitrage-free) model of the term structure. That said, the benefits may outweigh the cost for many practical applications.⁷ Indeed, a positive trade-off is suggested by the applications already mentioned above and my own empirical results presented later for the

to maturity, employs the simplification that the parameters used to estimate the state variables in the ZLB-constrained periods are obtained from a GATSM estimated with data from the non-ZLB-constrained period. Richard (2013) uses a lattice method and data out to 30 years time to maturity, but on p. 40 notes that “it requires a long time, literally a month, on large and fast computers to estimate”. Of course, once estimated parameters are available, subsequent implementations with those parameters will be very fast.

⁶See James and Webber (2000) for discussion on these points. For example, the standard deviation of Monte Carlo simulation results increases as the square root of time to maturity, and grid size and round-off issues can become material for finite difference methods with longer times to maturity.

⁷An analogy in this regard is that financial market practitioners, central bankers, and academics popular often apply the Nelson and Siegel (1987) model in its original (i.e. non-arbitrage-free) form, rather than using arbitrage-free Nelson and Siegel (1987) models or GATSMs. The implementation is more straightforward and the empirical results are similar. See Diebold and Rudebusch (2013) for further discussion.

United States; i.e. shadow short rates obtained with my framework are very similar to comparable results from a Black implementation. Furthermore, to show that my framework readily generalizes to more than two factors and longer maturity spans, I estimate two- and three-factor models with data sets out to 30 years time to maturity. I also establish that a robust method should be used for estimation, to appropriately allow for the inherent non-linearity of the ZLB-GATM framework, and that result may hold implications for the application of Black-GATSMs and the existing applications of my framework (or its equivalent) as noted above. In particular, I find that the extended Kalman filter gives very misleading estimates relative to those from an iterated extended Kalman filter.

The remainder of the article is arranged as follows. In section 2, I first outline the Black framework and its practical intuition, and then show how the Black framework may be re-expressed in terms of forward bond prices. The latter provides the link to section 3, where I show precisely how I approximate Black forward bond prices with a slightly alternative expression that leads to much more tractable results. Section 4 derives the expressions for my ZLB framework when GATSMs are used to represent the shadow short rate and the yield curve, which results in very straightforward closed-form analytic expressions for ZLB forward rates. Section 5 first shows how the shadow-GATSM state variable dynamics and the associated ZLB forward rate expressions may be used to form a state space representation for the ZLB term structure. It then presents the estimated shadow short rate results from various specifications and estimations, as mentioned earlier. Section 6 concludes.

2 The Black framework

In this section, I first summarize the foundations of the Black framework as they have already presented in the literature to date, and I also provide relevant intuition. I then re-express the Black framework in terms of Black forward bond prices, which provides the link to my proposed ZLB framework in section 3.

2.1 Black short rates and shadow/ZLB yield curve intuition

The fundamental building blocks for the Black framework are short rates $\underline{r}^B(t)$ that are constrained to be non-negative via the following mechanism:

$$\underline{r}^B(t) = \max\{0, r(t)\} \tag{1}$$

where $r(t)$ is the (default-free and risk-adjusted nominal) shadow short rate that can adopt negative values.⁸ In section 4, I will specify a GATSM process for $r(t)$, but at this stage I leave the exposition general to allow for any dynamic process. Note also that I use the underscore here and in all notation following to indicate quantities that are bounded below by zero, and I omit the underscore for shadow quantities.

⁸Black did not actually include the given expression, because it was a conceptual and descriptive paper. The applications of the Black framework referenced in section 1 all specify the expression while rightfully attributing it to Black. The name “shadow short rate” is from Black, but it does not necessarily imply a shadow price in the usual economic sense; i.e. it does not appear to have been obtained as a marginal change for a given objective function with respect to a constraint (although it may be possible to do so).

A given shadow short rate process subjected to the Black non-negative mechanism will produce a dynamic process for the Black short rate $\underline{r}^B(t)$ and, as I will outline in the following sub-section, the process for $\underline{r}^B(t)$ will in turn generate Black bond prices $\underline{P}^B(t, \tau)$, at time t and as a function of time to maturity τ . The Black yield curve is then the associated interest rate function $\underline{R}^B(t, \tau)$, which will be bounded below by zero at all times t and times to maturity τ . Note that a straightforward generalization at this stage would be to include a non-zero lower bound via $\underline{r}^B(t) = \max\{r_L, r(t)\}$, where r_L could be mildly positive or negative depending on institutional arrangements and market conditions.⁹ That said, I retain a zero lower bound for simplicity in notation and exposition, although I note later in section 4.4 how the end result of my framework could readily be adjusted to allow for a non-zero lower bound.

An alternative, and I believe more intuitive, expression of the Black ZLB framework is to explicitly decompose the Black short rate in equation 1 into its shadow short rate component and a component that represents the option value of physical currency contingent on the prevailing level of the shadow short rate. Specifically:

$$\underline{r}^B(t) = r(t) + \max\{-r(t), 0\} \quad (2)$$

With that decomposition, the Black yield curve at time t as a function of time-to-maturity τ may therefore be envisaged as a shadow yield curve $R(t, \tau)$ generated from the process for the shadow rate $r(t)$ plus a “physical currency option effect” $Z^B(t, \tau)$ generated from the process for $\max\{-r(t), 0\}$, i.e.:

$$\underline{R}^B(t, \tau) = R(t, \tau) + Z^B(t, \tau) \quad (3)$$

As I will formalize in the following sub-sections and section 3, $Z^B(t, \tau)$ reflects the cumulative value of a stream of European call options expiring between time t to time $t + \tau$. Those options give the investor the right to invest in physical currency (at an effective nominal interest rate of zero) at any point in time, against the alternative of being obliged to invest at the shadow short rate at all times, even if any realization should happen to be negative.

Figure 1 stylistically illustrates the decomposition of a Black yield curve into its shadow yield curve and option effect components. The shadow short rate is the value of the interest rate function at a time to maturity $\tau = 0$, i.e. $r(t) = R(t, 0)$, which is -5 percent in the example of figure 1. The associated option effect at $\tau = 0$ is $Z^B(t, 0) = \max\{-r(t), 0\} = \max\{-[-5 \text{ percent}], 0\} = 5$ percent, and the Black short rate $\underline{r}^B(t) = r(t) + \max\{-r(t), 0\} = -5 \text{ percent} + 5 \text{ percent} = 0$ percent.

⁹For example, the Bank of Japan’s policy operations have resulted in overnight and short-maturity rates prevailing at around positive 0.1 percent in the recent past, while the Denmark central bank deposit rate has been negative since July 2012. The German bond market has experienced instances of mildly negative interest rates even out to 2-year maturities, despite the positive policy rate setting by the European Central Bank, which likely results from the perceived safety of German bonds relative to other euro currency countries.

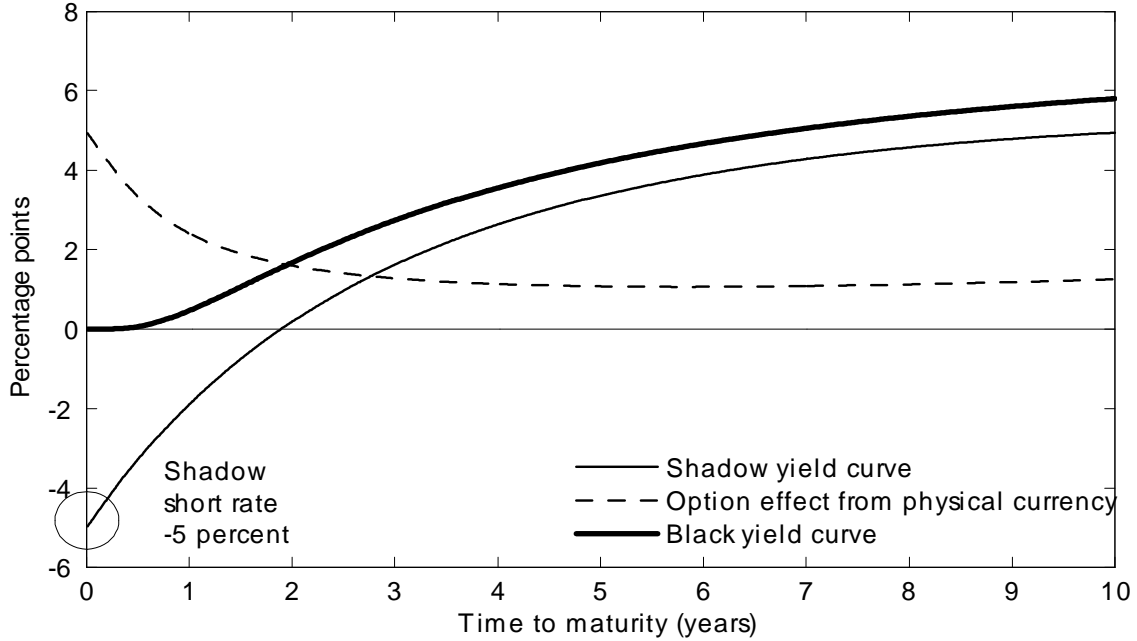


Figure 1: A stylized illustration of the Black yield curve decomposed into the shadow yield curve and the physical currency option effect.

2.2 Black bond prices

Black bond prices may be represented generically with the following standard arbitrage pricing expression:¹⁰

$$\underline{\mathbb{P}}^{\text{B}}(t, \tau) = \tilde{\mathbb{E}}_t \left\{ \exp \left(- \int_0^\tau \underline{r}^{\text{B}}(t+u) \, du \right) \cdot 1 \right\} \quad (4)$$

where 1 is the standard terminal cashflow for a bond (which I hereafter omit in similar expressions), $\underline{r}^{\text{B}}(t+u)$ is the risk-adjusted diffusion process for the Black short rate, and $\tilde{\mathbb{E}}_t$ is the risk-adjusted expectations operator based on information up to time t .

Equation 4 does not have closed-form analytic solution, as noted in section 1, but it can always be evaluated using numerical methods. For my exposition, I assume that Black bond prices are in principle obtained via a Monte Carlo simulation (hence the subscript “M”), and hence they can be expressed as follows:¹¹

$$\underline{\mathbb{P}}_{\text{M}}^{\text{B}}(t, \tau) = \frac{1}{J} \sum_{j=1}^J \underline{\mathbb{P}}_{\text{M},j}^{\text{B}}(t, \tau) \quad (5)$$

¹⁰See, for example Filipović (2009) p. 109. To facilitate the link with my later notation, I have adopted the Brace, Gatarek, and Musiela (1997) time and time-to-maturity notation for bond prices, as opposed to time and time-of-maturity notation. The latter would be $\underline{\mathbb{P}}^{\text{B}}(t, t+\tau) = \tilde{\mathbb{E}}_t \left\{ \exp \left(- \int_t^{t+\tau} \underline{r}(v) \, dv \right) \right\}$, the change of variables $u = v - t$ produces $\underline{\mathbb{P}}^{\text{B}}(t, t+\tau) = \tilde{\mathbb{E}}_t \left\{ \exp \left(- \int_0^\tau \underline{r}(t+u) \, du \right) \right\}$, and I denote $\underline{\mathbb{P}}^{\text{B}}(t, t+\tau)$ as $\underline{\mathbb{P}}^{\text{B}}(t, \tau)$.

¹¹See, for example, James and Webber (2000) p. 349. Ueno, Baba, and Sakurai (2006) and Bauer and Rudebusch (2013) are examples that use Monte Carlo simulations to implement the Black framework.

where $\underline{\mathbb{P}}_{M,j}^B(t, \tau)$ is a single simulated bond price:

$$\underline{\mathbb{P}}_{M,j}^B(t, \tau) = \exp \left[- \sum_{i=0}^I \underline{\mathbb{r}}_{M,j}^B(t + i \Delta t, \Delta t) \Delta t \right] \quad (6)$$

J represents an arbitrarily large number of simulations, and:

$$\underline{\mathbb{r}}_{M,j}^B(t + i \Delta t, \Delta t) = \max \{0, r_{M,j}(t + i \Delta t, \Delta t)\} \quad (7)$$

with $r_{M,j}(t + i \Delta t, \Delta t)$ a single simulated path of the discretized default-free shadow short rate diffusion process under the risk-adjusted measure. Regarding the discretization, $\Delta t = \tau/I$ (with I arbitrarily large to ensure that Δt is appropriately small), and for clarity in what follows I have used the notation $r_{M,j}(t + i \Delta t, \Delta t)$ to explicitly denote that the short rate prevailing at time $t + i \Delta t$ has a time to maturity of Δt (i.e. $r_{M,j}(t + i \Delta t, \Delta t)$ prevails over the period from $t + i \Delta t$ to $t + i \Delta t + \Delta t$).

2.3 Black forward bond prices

Black forward bond prices $\underline{\mathbb{P}}^B(t, \tau, \delta)$ at time t , for settlement (or delivery) at time $t + \tau$, and with a remaining time to maturity δ at the time of settlement may be expressed, using a standard term structure relationship,¹² as:

$$\underline{\mathbb{P}}_M^B(t, \tau, \delta) = \frac{\underline{\mathbb{P}}_M^B(t, \tau + \delta)}{\underline{\mathbb{P}}_M^B(t, \tau)} \quad (8)$$

where $\underline{\mathbb{P}}_M^B(t, \tau)$ has already been defined in the previous section, and:

$$\underline{\mathbb{P}}_M^B(t, \tau + \delta) = \frac{1}{J} \sum_{j=1}^J \underline{\mathbb{P}}_{M,j}^B(t, \tau + \delta) \quad (9)$$

with:

$$\underline{\mathbb{P}}_{M,j}^B(t, \tau + \delta) = \exp \left[- \sum_{i=0}^I \underline{\mathbb{r}}_{M,j}^B(t + i \Delta t, \Delta t) \Delta t \right] \cdot \exp [- \max \{0, r_{M,j}(t + \tau, \delta)\} \delta] \quad (10)$$

For use in the following section, I re-express the final term on the right-hand side of equation 10 as:

$$\begin{aligned} \exp [- \max \{ \cdot \} \delta] &= \exp [- \max \{0, r_{M,j}(t + \tau, \delta)\} \delta] \\ &= \exp [\min \{0, - r_{M,j}(t + \tau, \delta)\} \delta] \\ &= \min \{ \exp [0], \exp [- r_{M,j}(t + \tau, \delta)\} \delta] \} \\ &= \min \{1, \mathbb{P}_{M,j}(t + \tau, \delta)\} \\ &= \mathbb{P}_{M,j}(t + \tau, \delta) + \min \{1 - \mathbb{P}_{M,j}(t + \tau, \delta), 0\} \\ &= \mathbb{P}_{M,j}(t + \tau, \delta) - \max \{ \mathbb{P}_{M,j}(t + \tau, \delta) - 1, 0 \} \end{aligned} \quad (11)$$

¹²See, for example, Filipović (2009) p. 117.

and I use that result to express the Black forward bond price in equation 8 as:

$$\mathbb{P}_M^B(t, \tau, \delta) = \mathbb{P}_{M,1}^B(t, \tau, \delta) - \mathbb{P}_{M,2}^B(t, \tau, \delta) \quad (12)$$

where:

$$\mathbb{P}_{M,1}^B(t, \tau, \delta) = \frac{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I \mathbb{r}_{M,j}^B(t + i \Delta t) \Delta t \right] \cdot P_{M,j}(t + \tau, \delta)}{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I \mathbb{r}_{M,j}^B(t + i \Delta t) \Delta t \right]} \quad (13)$$

$$\mathbb{P}_{M,2}^B(t, \tau, \delta) = \frac{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I \mathbb{r}_{M,j}^B(t + i \Delta t) \Delta t \right] \cdot \max \{ P_{M,j}(t + \tau, \delta) - 1, 0 \}}{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I \mathbb{r}_{M,j}^B(t + i \Delta t) \Delta t \right]} \quad (14)$$

Having established this form for Black forward bond prices, I now proceed to approximate that expression in section 3.

3 Approximating the Black framework

In this section, I precisely define my proposed approximation to the Black forward bond prices introduced in section 2.3, and then show how that approximation leads to an analytic expression for discrete-period ZLB forward rates. Section 3.3 proceeds to transform the discrete-period ZLB forward rate expressions into continuous-time expressions and, in light of those expressions, section 3.4 revisits the intuition for the Black framework yield curve from 2.1.

3.1 Approximate Black forward bond prices

The approximation I make to the expression for Black forward bond prices is to discount the future cashflows $P_{M,j}(t + \tau, \delta)$ and $\max \{ P_{M,j}(t + \tau, \delta) - 1, 0 \}$ with shadow short rates instead of Black short rates. In notation, that means simply replacing the Black shadow short rates $\mathbb{r}_{M,j}^B(t + i \Delta t) \Delta t$ in equations 13 and 14 with shadow short rates $r_{M,j}(t + i \Delta t) \Delta t$. Equations 12, 13, and 14 therefore respectively become:

$$\mathbb{P}_M^C(t, \tau, \delta) = \mathbb{P}_{M,1}^C(t, \tau, \delta) + \mathbb{P}_{M,2}^C(t, \tau, \delta) \quad (15)$$

where:

$$\mathbb{P}_{M,1}^C(t, \tau, \delta) = \frac{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I r_{M,j}(t + i \Delta t) \Delta t \right] \cdot P_{M,j}(t + \tau, \delta)}{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I r_{M,j}(t + i \Delta t) \Delta t \right]} \quad (16)$$

and:

$$\mathbb{P}_{M,2}^C(t, \tau, \delta) = - \frac{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I r_{M,j}(t + i \Delta t) \Delta t \right] \cdot \max \{ P_{M,j}(t + \tau, \delta) - 1, 0 \}}{\frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I r_{M,j}(t + i \Delta t) \Delta t \right]} \quad (17)$$

I have used the notation $\underline{P}_M^C(t, \tau, \delta)$ to denote what I call the forward “currency-adjusted bond” price or forward CAB price. At this stage, the name and notation is simply to distinguish the quantity from the Black bond price $\underline{P}^B(t, \tau, \delta)$, but there is an appealing intuition underlying the name as I discuss below in section 3.2.

In light of the definition above, there are three immediate observations that can be made about the CAB framework in the context of ZLB yield curve modelling. First, the CAB framework is distinctly different from the Black framework given the differences in how future contingent cashflows are discounted in the two frameworks. Second and related, the CAB framework is clearly an approximate ZLB framework, unlike the Black framework. Specifically, the Black framework is arbitrage free and theoretically self-consistent because it uses own-model ZLB short rates for discounting future contingent cashflows, while the CAB framework does not. Third, the approximation of the CAB framework to the Black framework is likely to be better for shorter times to maturity and worse for longer times to maturity. The basis for that conjecture is that differences between the distribution of shadow short rates and ZLB short rates will increase by horizon/time to maturity, which should therefore result in larger differences between the discount factors in the two respective frameworks. Indeed, that pattern of approximation divergences with the Black framework by time to maturity has already been established in the related articles mentioned in the introduction, but the maximum reported differences are small. That is, the one-factor CAB-GATSM in Krippner (2013) has a maximum divergence of 13 basis points for the 30-year yield, the restricted three-factor CAB-GATSM in Christensen and Rudebusch (2013) has a maximum divergence of 4 basis points for the 10-year yield, and the three-factor GATSM of Wu and Xia (2013) has a maximum divergence of 0.92 basis points for the 10-year yield. These results are highly encouraging, although systematically investigating the divergences by model specification, parameter values, and times to maturity remains an important topic for future research.

3.2 Re-expressing of CAB forward bond prices

When the shadow rate process admits closed form analytic solutions in continuous time, which will be the case for generic GATSM specifications that I use in section 4 (and some other processes)¹³, then all of the individual summations underlying the forward CAB price in equation 15 can be expressed more succinctly. At this stage, I simply indicate that closed-form analytic solutions are available by returning to generic arbitrage pricing expressions obtained in the limit of $\Delta t \rightarrow 0$. Hence, the denominator

¹³For example, closed-form analytic solutions are available for Cox, Ingersoll, and Ross (1985)/square-root models with independent factor innovations (although the expressions are relatively complex; e.g. see Chen and Scott (1992) for the two-factor case). Therefore, the CAB framework could be used for mixed Gaussian/square-root models subject to the ZLB. Closed-form analytic solutions are also available for some jump diffusion models, depending on how the jumps are defined; see Piazzesi (2010) p. 709-17 for further discussion.

of both $\underline{P}_{M,1}^C(t, \tau, \delta)$ and $\underline{P}_{M,2}^C(t, \tau, \delta)$ is the shadow bond price $P(t, \tau)$:

$$\begin{aligned} P(t, \tau) &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I r_{M,j}(t + i \Delta t) \Delta t \right] \right\} \\ &= \tilde{\mathbb{E}}_t \left\{ \exp \left(- \int_0^\tau r(t + u) du \right) \right\} \end{aligned} \quad (18)$$

the numerator for $\underline{P}_{M,1}^C(t, \tau, \delta)$ is the shadow bond price $P(t, \tau + \delta)$:

$$\begin{aligned} P(t, \tau + \delta) &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I r_{M,j}(t + i \Delta t) \Delta t \right] \cdot P_{M,j}(t + \tau, \delta) \right\} \\ &= \tilde{\mathbb{E}}_t \left\{ \exp \left(- \int_0^{\tau+\delta} r(t + u) du \right) \right\} \end{aligned} \quad (19)$$

and the numerator for $\underline{P}_{M,2}^C(t, \tau, \delta)$ is the shadow bond option price $C(t, \tau, \delta)$:

$$\begin{aligned} C(t, \tau, \delta) &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{1}{J} \sum_{j=1}^J \exp \left[- \sum_{i=0}^I r_{M,j}(t + i \Delta t) \Delta t \right] \cdot \max \{ P_{M,j}(t + \tau, \delta) - 1, 0 \} \right\} \\ &= \tilde{\mathbb{E}}_t \left\{ \exp \left(- \int_0^{\tau+\delta} r(t + u) du \right) \right\} \cdot \max \{ P(t + \tau, \delta) - 1, 0 \} \end{aligned} \quad (20)$$

which is at time t , for expiry at time $t + \tau$, and with a remaining time to maturity δ on the bond at the time of the option expiry.

The forward CAB price $\underline{P}^C(t, \tau, \delta)$ may therefore be expressed as:

$$\begin{aligned} \underline{P}^C(t, \tau, \delta) &= \frac{P(t, \tau + \delta)}{P(t, \tau)} - \frac{C(t, \tau, \delta)}{P(t, \tau)} \\ &= P(t, \tau, \delta) - Z_P^C(t, \tau, \delta) \end{aligned} \quad (21)$$

where $Z_P^C(t, \tau, \delta)$ is the value of the option to hold physical currency in the CAB framework.

The name forward ‘‘currency-adjusted bond’’ price highlights that the fundamental building block for my ZLB framework is a forward shadow bond adjusted by a call option that represents the availability of physical currency as an alternative investment. Specifically, investors at time $t + \tau$ have the choice of purchasing the shadow bond at a price $P_j(t + \tau, \delta) = \exp[-r_j(t + \tau)\delta]$ if $P_j(t + \tau, \delta) < 1$ (hence locking in a return of $r_j(t + \tau) > 0$ over the time step from $t + \tau$ to $t + \tau + \delta$) or holding physical currency with a price of 1 which will retain its price of 1 into the next period (hence locking in a return $r_j(t + \tau) = 0$). The option of holding physical currency is equivalent to the investor possessing an option to purchase the shadow bond at a price $P_j(t + \tau, \delta) = 1$, given the known pay-off of 1 at time $t + \tau + \delta$ and the original assumption from section 2.2 that shadow short rates are default-free and under the risk-adjusted measure (which naturally translates to the associated single-step shadow bond).

Correspondingly, at time t , investors have the following marginal choices with respect to time $t + \tau$: (1) arrange to have 1 physical currency unit become available at

time $t + \tau$ by purchasing the CAB $P^C(t, \tau)$, which will mature with a pay-off of 1 at time $t + \tau$; or (2) commit at time t to extending the investment in the CAB $P^C(t, \tau)$ by purchasing $P^C(t, \tau + \delta) / P^C(t, \tau)$ units of the forward CAB $P^C(t, \tau, \delta)$, which effectively converts the CAB $P^C(t, \tau)$ into the CAB $P(t, \tau + \delta)$. The first marginal choice will provide the investor with the option to hold physical currency over the time step from $t + \tau$ to $t + \tau + \delta$, i.e. to invest at $P_j(t + \tau, \delta) = 1$ if $P_j(t + \tau, \delta) > 1$, thereby avoiding a negative return. The second marginal choice forgoes the latter option, but the appropriate amount of $Z_P^C(t, \tau, \delta)$ has already been embedded in the CAB price $P^C(t, \tau + \delta)$ to compensate the investor,¹⁴ thereby making them indifferent to either choice.

A useful and simple illustration of the intuition above is to set $\tau = 0$. In that case, the forward CAB price obviously embeds the value of the option to hold physical currency, because:

$$\begin{aligned}
\underline{P}^C(t, 0, \delta) &= \underline{P}^C(t, \delta) \\
&= \min\{1, P(t, \delta)\} \\
&= P(t, \delta) + \min\{1 - P(t, \delta), 0\} \\
&= P(t, \delta) - \max\{P(t, \delta) - 1, 0\} \\
&= P(t, \delta) - C(t, 0, \delta)
\end{aligned} \tag{22}$$

where $C(t, 0, \delta)$ is the known pay-off from an expiring call option on the shadow bond $P(t, \delta)$.

3.3 Continuous-time CAB term structure model

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 often more convenient to work with term structure models in continuous time. Furthermore, using forward rates is particularly convenient in the Gaussian context, as we will see in sections 4 and 5. Also, from this point onward, I omit the superscript ‘‘C’’ for CAB quantities, and so quantities with an underscore represent values constrained by the ZLB within the CAB framework. (I continue to use the superscript ‘‘B’’ for Black framework quantities.)

CAB forward rates $\underline{f}(t, \tau, \delta)$ at time t and over the discrete time-to-maturity step from τ to $\tau + \delta$ are given by the following standard relationship:¹⁵

$$\underline{f}(t, \tau, \delta) = -\frac{1}{\delta} \log [P(t, \tau, \delta)] \tag{23}$$

and those discrete-period forward rates can be converted into instantaneous forward rates $\underline{f}(t, \tau)$, at time t and as a function of horizon/time to maturity τ , by the following standard relationship:

$$\begin{aligned}
\underline{f}(t, \tau) &= \lim_{\delta \rightarrow 0} \{\underline{f}(t, \tau, \delta)\} \\
&= \lim_{\delta \rightarrow 0} \{P(t, \tau, \delta) - Z_P(t, \tau, \delta)\}
\end{aligned} \tag{24}$$

¹⁴The amount $Z_P^C(t, \tau, \delta)$ is appropriate in the context my approximate ZLB framework, but the Black framework compensation will be different, as I discuss in section 3.3. Also note that the borrower’s perspective is simply the direct counterpart to the investor’s perspective that I have used for the discussion.

¹⁵References for any standard term structure relationships I use in this section and subsequently in the article are, for example, Filipović (2009) p. 7 or James and Webber (2000) chapter 3.

Equation 24 can be simplified. Appendix A.1 contains the details of that derivation, and the final result is:

$$\underline{f}(t, \tau) = f(t, \tau) + z(t, \tau) \quad (25)$$

where $f(t, \tau)$ is the instantaneous shadow forward rate curve, and $z(t, \tau)$ is the instantaneous CAB option effect for the forward rate. At this stage, I express that result in its implicit form, i.e.:

$$z(t, \tau) = \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \delta)}{P(t, \tau)} \right] \right\} \quad (26)$$

but that implicit form will generate explicit closed-form analytic expressions whenever such expressions are available for bond prices $P(t, \tau)$ and option prices $C(t, \tau, \delta)$.

Having defined the CAB forward rate $\underline{f}(t, \tau)$, the CAB interest rate $\underline{R}(t, \tau)$ is readily defined using the standard continuous-time term structure relationships, i.e.:

$$\begin{aligned} \underline{R}(t, \tau) &= \frac{1}{\tau} \int_0^\tau \underline{f}(t, u) \, du \\ &= \frac{1}{\tau} \int_0^\tau f(t, u) \, du + \frac{1}{\tau} \int_0^\tau z(t, u) \, du \\ &= R(t, \tau) + Z(t, \tau) \end{aligned} \quad (27)$$

where u is a dummy integration variable for horizon/time to maturity, $R(t, \tau)$ is the shadow interest rate function, and $Z(t, \tau)$ is the CAB option effect for interest rates.

Similarly, $\underline{P}(t, \tau)$ may be expressed as the product of the shadow bond price $P(t, \tau)$ and an option effect for $Z_P(t, \tau)$, i.e.:

$$\begin{aligned} \underline{P}(t, \tau) &= \exp \left[- \int_0^\tau \underline{f}(t, u) \, du \right] \\ &= \exp [-\tau \cdot \underline{R}(t, \tau)] \\ &= \exp [-\tau \cdot R(t, \tau)] \cdot \exp [-\tau \cdot Z(t, \tau)] \\ &= P(t, \tau) \cdot Z_P(t, \tau) \end{aligned} \quad (28)$$

where:

$$Z_P(t, \tau) = \exp [-\tau \cdot Z(t, \tau)] \quad (29)$$

The instantaneous CAB short rate $\underline{r}(t)$ associated with any continuous-time dynamic process for the shadow short rate $r(t)$ is a well-defined and non-negative quantity. Appendix A.2 contains the details, which I summarize as follows:

$$\begin{aligned} \underline{r}(t) &= \lim_{\tau \rightarrow 0} \underline{f}(t, \tau) \\ &= f(t, 0) + z(t, 0) \\ &= r(t) + \max \{-r(t), 0\} \\ &= \max \{r(t), 0\} \end{aligned} \quad (30)$$

3.4 The Black yield curve decomposition

Having derived the shadow yield curve and option effects for the CAB framework, I now briefly return to formalize the intuition of the Black framework from section 2.1.

Using the currency-adjustment intuition developed above for the CAB framework, the Black yield curve may be decomposed into the shadow yield curve and an option effect for interest rates as follows:

$$\begin{aligned}
\mathbb{R}^B(t, \tau) &= \frac{1}{\tau} \int_0^\tau \mathbb{f}^B(t, u) \, du \\
&= \frac{1}{\tau} \int_0^\tau f(t, u) \, du + \frac{1}{\tau} \int_0^\tau z^B(t, u) \, du \\
&= \mathbb{R}(t, \tau) + Z^B(t, \tau)
\end{aligned} \tag{31}$$

Unlike the CAB framework, the Black interest rate expression $\mathbb{R}^B(t, \tau)$ is theoretically self-consistent, for the reasons previously discussed in section 3.1. Of course, the relative disadvantage is that $Z^B(t, \tau)$ cannot be obtained using closed-form analytic solutions that may be available for some specifications of the shadow term structure, such as the GATSM specification to which I now turn.

4 The CAB-GATSM framework

In this section, I develop the generic CAB-GATSM framework using the continuous-time CAB framework established in section 3.3 and a generic GATSM specification to represent the shadow term structure. Section 4.1 outlines the GATSM specification, and sections 4.2 and 4.3 respectively derive the shadow-GATSM forward rate and the CAB-GATSM option effect that together provide the closed-form analytic expression for CAB-GATSM forward rates.

4.1 Shadow-GATSM short rate process

I adopt the generic GATSM specification from Dai and Singleton (2002) pp. 437-38 as a convenient point of reference, for the compactness of its matrix notation and because it already contains explicit bond price solutions for most GATSM specifications. Hence, the generic shadow-GATSM short rate process is:

$$r(t) = a_0 + b_0' x(t) \tag{32}$$

where a_0 is a constant, b_0 is a constant $N \times 1$ vector containing the weights for the N state variables $x_n(t)$, and $x(t)$ is an $N \times 1$ vector containing the N state variables $x_n(t)$. Under the physical or \mathbb{P} measure, $x(t)$ evolves as the following correlated vector Ornstein-Uhlenbeck process:

$$dx(t) = \kappa [\theta - x(t)] \, dt + \sigma dW(t) \tag{33}$$

where θ is a constant $N \times 1$ vector representing the long-run level of $x(t)$, κ is a constant $N \times N$ matrix that governs the deterministic mean reversion of $x(t)$ to θ , σ is a constant $N \times N$ matrix representing the potentially correlated volatilities of $x(t)$, and $dW(t)$ is an $N \times 1$ vector with independent Wiener components $dW_n(t) \sim N(0, 1) \sqrt{dt}$.

The market prices of risk are linear with respect to the state variables, i.e.:¹⁶

$$\Pi(t) = \sigma^{-1} [\gamma + \Gamma x(t)] \tag{34}$$

¹⁶This is the “essentially affine” specification from Duffee (2002), but for a model with full Gaussian dynamics. Also see Cheridito, Filipović, and Kimmel (2007) for further discussion on market price of risk specifications.

where γ and Γ are respectively a constant $N \times 1$ vector and constant $N \times N$ matrix. The risk-adjusted process for $x(t)$ is:

$$dx(t) = \tilde{\kappa} [\tilde{\theta} - x(t)] dt + \sigma dW(t) \quad (35)$$

where $\tilde{\kappa} = \kappa + \Gamma$ and $\tilde{\theta} = \tilde{\kappa}^{-1}(\kappa\theta - \gamma)$.

4.2 Shadow-GATSM forward rates

I derive the shadow-GATSM forward rate expression $f(t, \tau)$ by solving for GATSM bond prices and then applying the standard term structure expression relating bond prices to forward rates. The details are contained in appendix B, but to summarize, the shadow-GATSM bond prices $P(t, \tau)$ I obtain have the standard exponential affine form, i.e.:

$$P(t, \tau) = \exp[-A(\tau) - B(\tau)'x(t)] \quad (36)$$

where:

$$B(\tau) = [\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]' x(t) \quad (37)$$

and $[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]$ is the following $N \times 1$ vector function of forward rate loadings:

$$[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)] = b_0' \exp(-\tilde{\kappa}\tau) \quad (38)$$

For the purposes of the exposition here, $A(\tau)$ is most conveniently defined via its derivative, i.e.:¹⁷

$$\frac{d}{d\tau}A(\tau) = a_0 + \int_0^\tau [\mathfrak{h}(b_0, \tilde{\kappa}, v-s)]' \sigma' \left[\int_s^\tau \sigma [\mathfrak{h}(b_0, \tilde{\kappa}, u-s)] du \right] ds \quad (39)$$

In appendix B.2, I apply the following standard term structure expression to equation 36 to obtain the given expression for forward rates:

$$\begin{aligned} f(t, \tau) &= -\frac{d}{d\tau} \log [P(t, \tau)] \\ &= A(\tau) + B(\tau)'x(t) \\ &= a_0 + [\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]' x(t) \\ &\quad + \int_0^\tau [\mathfrak{h}(b_0, \tilde{\kappa}, \tau-s)]' \sigma' \left[\int_s^\tau \sigma [\mathfrak{h}(b_0, \tilde{\kappa}, u-s)] du \right] ds \end{aligned} \quad (40)$$

From here on, I use an intuitive interpretation of equation 40 as a volatility effect and the expected path of the short rate, as at time t , under the risk-adjusted measure. That is:

$$f(t, \tau) = \tilde{\mathbb{E}}_t [r(t + \tau)] + \text{VE}(\tau) \quad (41)$$

where:

$$\tilde{\mathbb{E}}_t [r(t + \tau)] = a_0 + [\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]' x(t) \quad (42)$$

¹⁷Dai and Singleton (2002) provides an expression for $A(\tau)$ that applies in the case of distinct eigenvalues for $\tilde{\kappa}$, but my result is generalized to allow for any GATSM. In particular, it accommodates shadow-GATSM specifications with repeated eigenvalues for $\tilde{\kappa}$ which is one of the models that I later derive and estimate in section 5.

and the volatility effect is $VE(\tau) = \int_0^\tau [\mathbb{1}(b_0, \tilde{\kappa}, \tau - s)]' \sigma' \left[\int_s^\tau \sigma [\mathbb{1}(b_0, \tilde{\kappa}, u - s)] du \right] ds$.¹⁸

Note that the functional expression for $\tilde{\mathbb{E}}_t[r(t + \tau)]$ by horizon/time to maturity τ could be used to obtain the expected time to a given value of the shadow short rate conditional on the current state variable vector and the model specification. For example, if the current value of the shadow short rate $r(t) = a_0 + b_0' x(t)$ is negative, then the expected time to the “lift-off” value of zero could be obtained by setting $\tilde{\mathbb{E}}_t[r(t + \tau)]$ to zero and solving for τ_0 , i.e.:

$$0 = a_0 + [\mathbb{1}(b_0, \tilde{\kappa}, \tau_0)]' x(t) \quad (43)$$

4.3 CAB-GATSM option effect

The CAB-GATSM option effect $z(t, \tau)$ may be derived using the limit expression for the option effect from equation 26 and closed-form analytic expressions for the prices of bonds and call options for GATSMs, as I show that explicitly in appendix C. However, it turns out that using GATSMs to represent the shadow term structure allows a much more direct derivation.¹⁹ The details are contained in appendix D, but I summarize them here as follows.

First I establish that, under the risk-adjusted $t + \tau$ forward measure, the distribution of risk-adjusted shadow-GATSM short rates for the generic shadow-GATSM has the following normal distribution:

$$r(t + \tau) \sim N \left\{ f(t, \tau), [\omega(\tau)]^2 \right\} \quad (44)$$

where:

$$\begin{aligned} [\omega(\tau)]^2 &= \text{var}[r(t + \tau)] \\ &= b_0' \text{cov}[x(t + \tau)] b_0 \\ &= b_0' \left[\int_0^\tau [\exp(-\tilde{\kappa}s)]' \sigma' \sigma [\exp(-\tilde{\kappa}s)] ds \right] b_0 \\ &= \int_0^\tau [\mathbb{1}(b_0, \tilde{\kappa}, s)]' \sigma' \sigma [\mathbb{1}(b_0, \tilde{\kappa}, s)] ds \end{aligned} \quad (45)$$

The CAB-GATSM option effect is then the expected value of $\max\{-r(t + \tau), 0\}$

¹⁸I refer to $VE(\tau)$ as the volatility effect because it captures the influence that volatility in the shadow short rate has on expected returns due to Jensen’s inequality. That is, the expected compounded return from investing in a volatile short rate over time t to $t + \tau$ is less than the compounded return from investing in the expected short rate over the same period.

¹⁹I am indebted to Scott Richard for pointing out that my original result in appendix C may be obtained as the expected value of a normal distribution truncated at zero, which led me to the forward measure derivation. I have chosen to value the option effect $z(t, \tau)$ here, which aligns with my original exposition, but $f(t, \tau)$ could be calculated directly as in equation 50 with $r_L = 0$. Also note that there is no need to explicitly calculate GATSM bond option expressions to evaluate the option effect expression for any CAB-GATSM, as suggested by Christensen and Rudebusch (2013) p. 9; one should just use the GATSM volatility and forward rate expressions directly.

under the same risk-adjusted $t + \tau$ forward measure, i.e.:

$$\begin{aligned}
z(t, \tau) &= \tilde{\mathbb{E}}_{t+\tau} [\max \{-r(t + \tau), 0\}] \\
&= \int_{-\infty}^{\infty} \max \{-r(t + \tau), 0\} \cdot \text{Prob}[r(t + \tau)] \, d r(t + \tau) \\
&= \int_{-\infty}^0 -r(t + \tau) \cdot \text{Prob}[r(t + \tau)] \, d r(t + \tau)
\end{aligned} \tag{46}$$

where the integral from 0 to ∞ is set to zero because $\max \{-r(t + \tau), 0\} = 0$ over that domain.

Given the normal distribution in equation 44, the probability density function for $r(t + \tau)$ is:

$$\text{Prob}[r(t + \tau)] = \frac{1}{\omega(\tau)} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{r(t + \tau) - f(t, \tau)}{\omega(\tau)}\right]^2\right) \tag{47}$$

and so $\underline{f}(t, \tau)$ may be evaluated directly as:

$$\begin{aligned}
z(t, \tau) &= \int_0^{\infty} -r(t + \tau) \cdot \frac{1}{\omega(\tau)} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{r(t + \tau) - f(t, \tau)}{\omega(\tau)}\right]^2\right) \, d[r(t + \tau)] \\
&= -f(t, \tau) \left(1 - \Phi\left[\frac{f(t, \tau)}{\omega(\tau)}\right]\right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right)
\end{aligned} \tag{48}$$

where $\Phi[\cdot]$ is the cumulative normal distribution function.

4.4 CAB-GATSM forward rates

Substituting the result for $z(t, \tau)$ from equation 94 into equation 25 gives the generic CAB-GATSM forward rate expression:

$$\begin{aligned}
\underline{f}(t, \tau) &= f(t, \tau) + z(t, \tau) \\
&= f(t, \tau) - f(t, \tau) \left(1 - \Phi\left[\frac{f(t, \tau)}{\omega(\tau)}\right]\right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right) \\
&= f(t, \tau) \cdot \Phi\left[\frac{f(t, \tau)}{\omega(\tau)}\right] + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right)
\end{aligned} \tag{49}$$

where explicit generic expressions are given for $f(t, \tau)$ and $\omega(\tau)$ respectively in equations 40 and 45.²⁰

Two observations immediately follow from the result above. First, CAB-GATSM short rates are Markovian; i.e. they have no “memory” or path dependency, because

²⁰Equation 49 is equivalent to Wu and Xia (2013) eq. 6 with a lower bound of zero. That is, my $f(t, \tau)$ and $\omega(\tau)$ expressions are respectively the continuous-time analogs of their $f_{n,n+1,t}^{GA}$ and $(\sigma_n^Q)^2$, my $f(t, \tau) / \omega(\tau)$ is their z , and my $\omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right)$ is their normal density function $\phi(z)$. I note how to incorporate a non-zero lower bound in my framework in equation 50 immediately below.

the distribution of shadow short rates is only conditional on the initial state and the horizon/time to maturity. Conversely, Black-GATSM short rates are non-Markovian.²¹ The second observation is that incorporating a non-zero lower bound r_L into the CAB-GATSM framework becomes very straightforward, i.e.:

$$\underline{f}(t, \tau) = \int_{r_L}^{\infty} r(t + \tau) \cdot \text{Prob}[r(t + \tau)] \, dr(t + \tau) \quad (50)$$

ZLB-GATSM interest rates $\underline{R}(t, \tau)$ and bond prices $\underline{P}(t, \tau)$ may be obtained using the results for $\underline{f}(t, \tau)$ within the respective expressions provided in equations 27 and 28. For example, I use the interest rate expression as the basis for my empirical estimations in section 5.

4.5 Observations on the CAB-GATSM framework

One advantage of developing the CAB-GATSM framework based on a generic GATSM specification for the shadow term structure is that the associated observations will apply to any particular CAB-GATSM specification, irrespective of the number of factors and factor inter-relationships. Bearing that generality in mind, I make the five observation about the CAB-GATSM framework.

First, the CAB-GATSM framework obviously preserves the complete flexibility of the GATSM class of models by using the GATSM as its foundation. The CAB-GATSM framework provides the explicit modification $z(t, \tau)$ to ensure that the ZLB is respected by the entire CAB-GATSM forward rate curve at each point in time, no matter how much shadow-GATSM forward rates evolve below zero. Specifically, as $f(t, \tau)$ decreases to larger negative values, $\Phi[f(t, \tau)/\omega(\tau)]$ and $\exp(-\frac{1}{2}[f(t, \tau)/\omega(\tau)]^2)$ both approach zero, so $\lim_{f(t, \tau)/\omega(\tau) \rightarrow -\infty} \underline{f}(t, \tau) = 0$.

The second observation is that the CAB-GATSM forward rate curve will always be a simple closed-form analytic expression. That is evident from the generic CAB-GATSM being itself composed of simple closed-form analytic expressions, i.e.: (1) $f(t, \tau)$, which is defined by the state variable vector $x(t)$ and scalar exponential functions of time to maturity τ ; (2) $\omega(\tau)$, which is defined by state variable innovation variances and covariances and scalar exponential functions of τ ; (3) the standard cumulative normal function $\Phi[f(t, \tau)/\omega(\tau)]$; and (4) the scalar exponential function $\exp(-\frac{1}{2}[f(t, \tau)/\omega(\tau)]^2)$.

Third, CAB-GATSM interest rates $\underline{R}(t, \tau)$ for any given time to maturity τ must be obtained via numerical integration, a property that arises from the Gaussian context.²² However, the number of numerical evaluations required for such integrals remains invariant to the specification of the CAB-GATSM because $\underline{f}(t, \tau)$ is always a closed-form analytic expression; i.e. the evaluation of $\underline{R}(t, \tau)$ is always an elementary univariate numerical integral of $\underline{f}(t, \tau)$ only with respect to τ . If CAB-GATSM bond prices $\underline{P}(t, \tau)$ are required, they will be the scalar exponential of the same (negated) univariate numerical integral.

²¹As a corollary, the Heath, Jarrow, and Morton (1992) framework and methods introduced for that literature may offer convenient and efficient methods of applying the Black framework, which would be an interesting topic to follow up in future research.

²²I have tried, obviously without success, to derive closed-form analytic expressions for $\underline{r}(t, \tau)$, and hence $\underline{P}(t, \tau)$. The relatively simple form of $\underline{f}(t, \tau)$ tantalizingly suggests the possibility of an analytic integral, but integration by parts was unsuccessful as was “brute force” analytic integration via Mathematica.

Fourth, the very tractable evaluation of CAB-GATSM interest rates (or bond prices) along with multivariate-normal transition densities for the CAB-GATSM state variables means that CAB-GATSMs will retain a large degree of the GATSM-class tractability for empirical applications and estimations.

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

5 Estimating the CAB-GATSM framework

In this section, I use the iterated extended Kalman filter (hereafter IEKF) to estimate a variety of CAB-GATSMs with different sets of United States yield curve data.²³ Section 5.1 first introduces the state space representation common to all CAB-GATSM specifications. Sections 5.2, 5.3, and 5.4 then respectively specify three different CAB-GATSMs and present the results from their estimation. In section 5.5, I summarize the key results and make some general observations.

5.1 CAB-GATSM state space representation

The state equation for the CAB-GATSM is the state equation for the GATSM used to represent the shadow term structure, which may in turn be obtained from the conditional expectation of $x(t + \tau)$ under the physical measure, i.e. $\mathbb{E}_t\{x(t + \tau) | x(t)\}$, with discrete time steps Δt for τ . The result is standard in the literature, e.g. see Babbs and Nowman (1998), but I also provide a derivation in appendix E.2.2. Hence, the state equation for any CAB-GATSM is always a first-order vector autoregression for the state variables, i.e.:

$$x_{t+1} = [I - F]\theta + Fx_t + \varepsilon_{t+1} \quad (51)$$

where x_t is the $N \times 1$ vector of state variables (with t as a subscript being an integer index from 1 to T representing the discrete time intervals Δt at which data is observed over the sample period), $F = \exp(-\kappa\Delta t)$ is a constant $N \times N$ matrix, $[I - F]\theta$ is a constant $N \times 1$ vector, and ε_t is an $N \times 1$ vector of innovations with a constant $N \times N$ covariance matrix Ω_ε .

The measurement equation for the CAB-GATSM is:

$$\begin{bmatrix} \mathbb{R}_t(\tau_1) \\ \vdots \\ \mathbb{R}_t(\tau_K) \end{bmatrix} = \begin{bmatrix} \mathbb{R}(x_t, \mathbb{A}, \tau_1) \\ \vdots \\ \mathbb{R}(x_t, \mathbb{A}, \tau_K) \end{bmatrix} + \begin{bmatrix} \eta_t(\tau_1) \\ \vdots \\ \eta_t(\tau_K) \end{bmatrix} \quad (52)$$

where $\mathbb{R}_t(\tau_k)$ is the interest rate observed at t for time-to-maturity τ_k ; $\mathbb{R}(x_t, \mathbb{A}, \tau_k)$ is the CAB-GATSM interest rate as a function of the state variables x_t , the parameter

²³Details on the IEKF as relevant to my application are contained in appendix E of the working paper version of this article.

set \mathbb{A} for the given model, and the time-to-maturity τ_k ; and $\eta_t(\tau_k)$ is the component of $\mathbf{R}_t(\tau_k)$ that is unexplained by $\mathbf{R}(x_t, \mathbb{A}, \tau_k)$.

CAB-GATSM interest rates for a given time to maturity τ_k are obtained by univariate numerical integration of the CAB-GATSM forward rates $\underline{f}(x_t, \mathbb{A}, i\Delta\tau)$.²⁴ Rectangular numerical integration using end-increment function evaluations, as I outline in appendix E.2.3, conveniently simplifies to an arithmetic mean, i.e.:

$$\mathbf{R}(x_t, \mathbb{A}, \tau_k) = \text{mean} \{ \underline{f}(x_t, \mathbb{A}, \Delta\tau), \underline{f}(x_t, \mathbb{A}, 2\Delta\tau), \dots, \underline{f}(x_t, \mathbb{A}, I_k\Delta\tau) \} \quad (53)$$

where $I_k = \tau_k/\Delta\tau$, and $\underline{f}(x_t, \mathbb{A}, \Delta\tau), \underline{f}(x_t, \mathbb{A}, 2\Delta\tau), \dots, \underline{f}(x_t, \mathbb{A}, I_k\Delta\tau)$ is a sequence of forward rates $\underline{f}(x_t, \mathbb{A}, u)$ for times to maturity $u = i\Delta\tau$.²⁵ That sequence is calculated from equation 49 using the sequences of shadow forward rates $f(x_t, \mathbb{A}, u)$ and annualized option volatilities $\omega(\mathbb{A}, u)$ for the given CAB-GATSM specification. Nevertheless, for any CAB-GATSM, the measurement equation may always be represented more generally and compactly in vector form, i.e.:

$$\mathbf{R}_t = \mathbf{R}(x_t, \mathbb{A}) + \eta_t \quad (54)$$

where \mathbf{R}_t is the $K \times 1$ vector representing the observed yield curve at t (i.e. the K interest rates of different times to maturity used to represent the yield curve at t), $\mathbf{R}(x_t, \mathbb{A})$ is the $K \times 1$ vector of CAB-GATSM interest rates $\mathbf{R}(x_t, \mathbb{A}, \tau_k)$, and η_t is the $K \times 1$ vector of unexplained yield curve components. I assume the $K \times K$ covariance matrix for the measurement equation is constant with the form $\Omega_\eta = \text{diag}[\sigma_\eta^2(\tau_1), \dots, \sigma_\eta^2(\tau_K)]$. That assumption is standard in the literature, and I also adopt the standard assumptions of zero intertemporal covariances and zero covariances between η_t and ε_t .

The key aspect to note about the measurement equation is that, unlike GATSMs, it involves non-linear functions of the state variables x_t . Therefore, estimating the state space representation requires a non-linear technique. I use the iterated extended Kalman filter (hereafter IEKF) for my estimations after finding that estimations using the extended Kalman filter (EKF) were not robust, which I discuss in section 5.3. The IEKF is covered in textbooks such as Simon (2006) p. 410-12 and Grewal and Andrews (2008) p. 312-13, and appendix E contains the details of the IEKF as relevant to my application. As noted in appendix E.1, I using the Nelder-Mead algorithm to obtain the parameter set \mathbb{A} and state variables x_t that maximize the log-likelihood function. Due to space constraints, I usually present only the shadow short rate results for each estimation; the full set of results for each estimation are available from the author on request, as as the results from estimations I mention but don't present.

5.2 CAB-GATSM(2)

In this section I develop the two-factor CAB-GATSM, or CAB-GATSM(2), by using the GATSM(2) to represent the shadow term structure within the CAB-GATSM framework. The maximally identifiable GATSM(2) has 12 free parameters, and I choose the following parameters to define the model: a_0 (one parameter), $\tilde{\kappa} = \text{diag}[\tilde{\kappa}_1, \tilde{\kappa}_2]$ (two

²⁴ Wu and Xia (2013) uses forward rates directly as observables, and so their measurement equation is left in forward rate terms.

²⁵ I set $\Delta\tau$ to 0.01 years, and found no apparent sensitivity in the results by using values of 0.1 or 0.001 years. I also found no sensitivity to the alternative numerical integration methods I tested.

parameters), κ (four parameters, because $\kappa = \tilde{\kappa} - \Gamma$, so Γ contributes four free parameters), θ (two parameters), and σ defined with the three parameters σ_1 , σ_2 , and ρ_{12} , i.e.:

$$\sigma' = \begin{bmatrix} \sigma_1 & 0 \\ \rho_{12}\sigma_2 & \sigma_2\sqrt{1 - \rho_{12}^2} \end{bmatrix} \quad (55)$$

Regarding the remaining parameters, I set $\tilde{\theta} = 0$, the non-diagonal elements of $\tilde{\kappa}$ to zero, and $b_0 = [1, 1]'$. Appendix F details the derivation of the shadow-GATSM(2) forward rate expression, with the final result:

$$\begin{aligned} f(x_t, \mathbb{A}, u) &= a_0 + x_{1,t} \cdot \exp(-\tilde{\kappa}_1 u) + x_{2,t} \cdot \exp(-\tilde{\kappa}_2 u) \\ &\quad - \sigma_1^2 \cdot \frac{1}{2} [G(\tilde{\kappa}_1, u)]^2 - \sigma_2^2 \cdot \frac{1}{2} [G(\tilde{\kappa}_2, u)]^2 \\ &\quad - \rho\sigma_1\sigma_2 \cdot G(\tilde{\kappa}_1, u) \cdot G(\tilde{\kappa}_2, u) \end{aligned} \quad (56)$$

where:

$$G(\phi, u) = \frac{1}{\phi} [1 - \exp(-\phi\tau)] \quad (57)$$

Note that $G(\phi, u)$ is a function that occurs extensively in GATSMs, and I use it hereafter where possible to keep expressions manageable.

To briefly provide some intuition of the shadow-GATSM(2) forward rate specification, note that the shadow short rate (hereafter SSR) at time t is defined as:

$$\begin{aligned} r_t &= a_0 + b_0' x_t \\ &= a_0 + x_{1,t} + x_{2,t} \end{aligned} \quad (58)$$

and the first line of equation 56 represents the expected path of the SSR as a function of horizon/time to maturity τ , i.e.:

$$\tilde{\mathbb{E}}_t [r(t + \tau)] = a_0 + x_{1,t} \cdot \exp(-\tilde{\kappa}_1\tau) + x_{2,t} \cdot \exp(-\tilde{\kappa}_2\tau) \quad (59)$$

where a_0 is the long-run (infinite-horizon) level of the shadow short rate, the state variable $x_{1,t}$ represents a latent component that captures more-persistent expected deviations of the SSR from a_0 , and the state variable $x_{2,t}$ captures less-persistent (i.e. faster mean-reverting) expected deviations of the SSR from a_0 . The remainder of equation 56 is the time-invariant volatility effect $\text{VE}(\tau)$.

Appendix F details the derivation of annualized option volatility for the CAB-GATSM(2), with the final result:

$$\omega(\mathbb{A}, u) = \sqrt{\sigma_1^2 \cdot G(2\tilde{\kappa}_1, u) + \sigma_2^2 \cdot G(2\tilde{\kappa}_2, u) + 2\rho\sigma_1\sigma_2 \cdot G(\tilde{\kappa}_1 + \tilde{\kappa}_2, u)} \quad (60)$$

Using these results in conjunction with the state space representation already given in section 5.1 completes the CAB-GATSM(2) specification. My first empirical application of the CAB-GATSM(2) is to the data described in Ichiue and Ueno (2013), i.e. month-end federal funds rates and 1, 2, 5, and 10-year rates from the Gurkaynak, Sack, and Wright (2007, hereafter GSW) data set for the period from January 1990 to March

2013.²⁶ I estimate the CAB-GATSM(2) with what I denote the IEKF(1e-8), which is the IEKF iterated until all elements of $|x_{t,i+1}^+ - x_{t,i+1}|$ are less than 1e-8, and obtain the estimated SSR series plotted in figure 2.²⁷ The correspondence with estimated Black-GATSM(2) SSRs from Ichiue and Ueno (2013) is close.²⁸

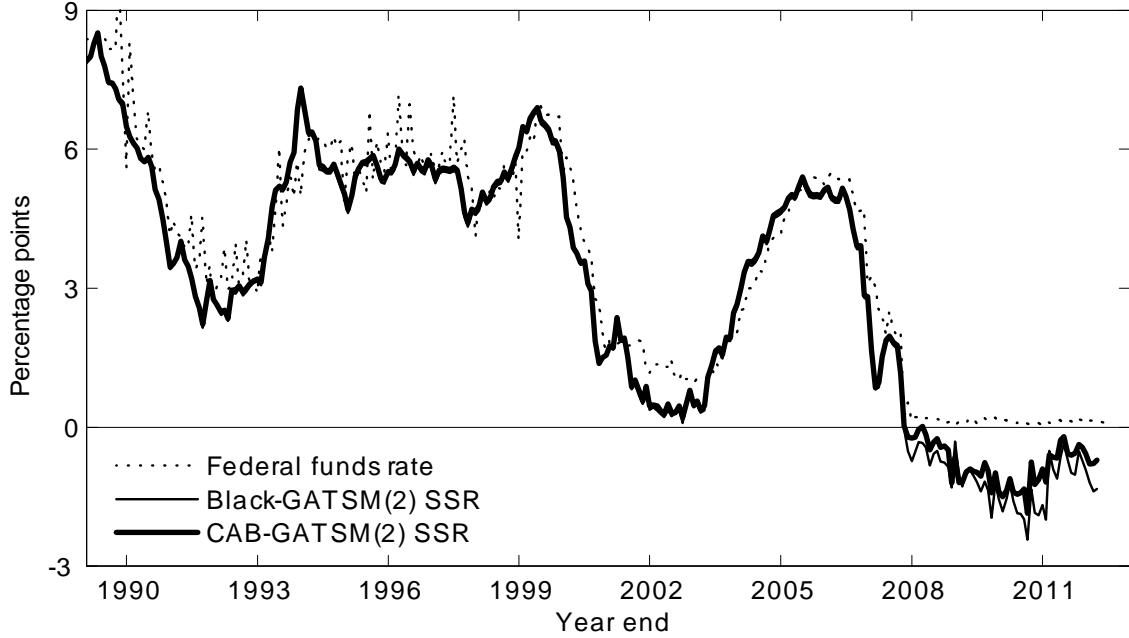


Figure 2: CAB-GATSM(2) shadow short rates estimated with the data set described in Ichiue and Ueno (2013), and Black-GATSM(2) shadow short rate estimates from Ichiue and Ueno (2013).

I also estimate the CAB-GATSM(2) with two alternative data sets over the same sample period, i.e.: (1) month-end 3- and 6-month Treasury bill rates and the 1, 2, 3, 4, 5, 7, and 10-year GSW rates, which I hereafter refer to as the GSW10 data set; and (2) the GSW10 data set with the addition of 20 and 30-year GSW rates, which I hereafter refer to as the GSW30 data set.²⁹ Using these two data sets allows a comparison of results obtained with different maturity spans, and estimations with the GSW30 data set illustrate that the CAB-GATSM framework may readily be applied to data with long maturity spans.

²⁶The GSW data set is maintained by the Federal Reserve Board and is available at <http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html>. I also allow for the constant 14 basis point lower bound imposed by Ichiue and Ueno (2013) from November 2009, although I found that it made little difference compared to results using a ZLB. I have converted the federal funds rates from the 360-day basis to a continuously compounding basis to match the basis of the GSW data, using a time to maturity of 1/365 years.

²⁷The value 1e-8 is 1/10000th of a basis point, so my choice of tolerance may be overly conservative and it could likely be relaxed in future empirical applications. However, for the purposes of the present article, I wanted to ensure that the state variables estimates had unambiguously converged at each point in time, and therefore were not propagating any material imprecision into subsequent state variable estimates.

²⁸I thank Yoichi Ueno for making the results available to me.

²⁹I have converted the Treasury bill rates from the 360-day discount basis to a continuously compounding basis to match the basis of the GSW data, using times to maturity of 0.25 and 0.5 years.

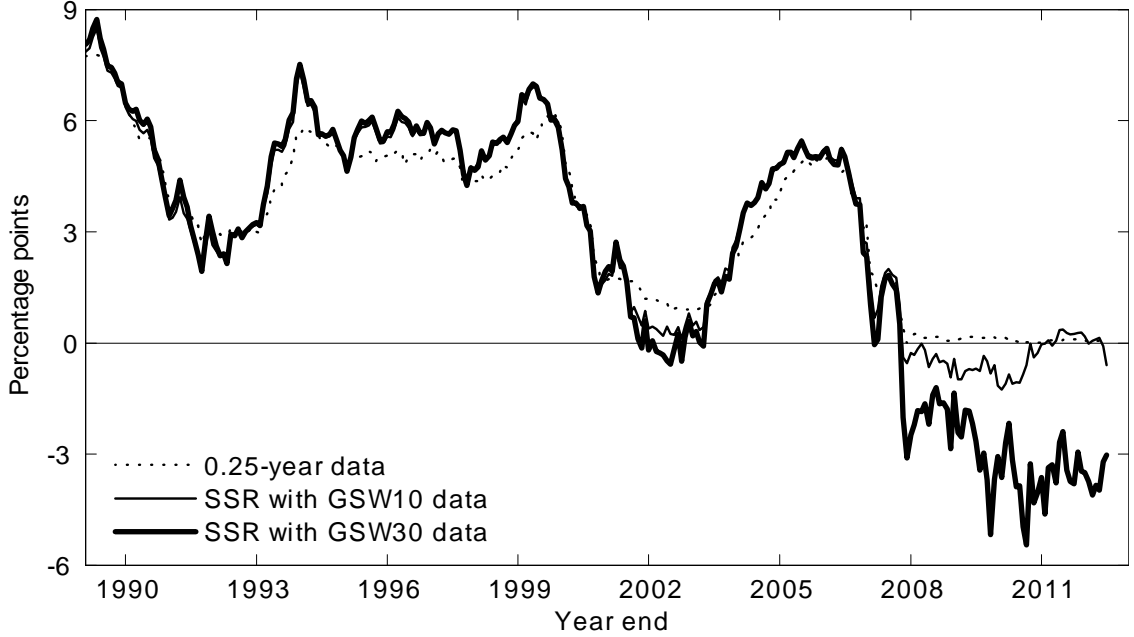


Figure 3: CAB-GATSM(2) shadow short rates estimated using the GSW10 and the GSW30 data sets and the IEKF(1e-8).

The results are plotted in figure 3 along with the shortest-maturity interest rate from both data sets (i.e. the 0.25-year rate). One immediate observation is that the estimated SSRs from the GSW10 and GSW30 data sets are distinctly different after late-2008 when interest rates became materially constrained by the ZLB. The SSR results obtained from the GSW10 data set also differ, albeit less substantially, from those obtained with the Ichiue and Ueno (2013) data set as plotted in figure 2.

5.3 Restricted CAB-GATSM(2)

The CAB-GATSM(2) can be made more parsimonious by setting the mean-reversion rate for the persistent component to zero, i.e. $\tilde{\kappa}_1 = 0$ in $\tilde{\kappa} = \text{diag}[\tilde{\kappa}_1, \tilde{\kappa}_2]$. The process for $x_{1,t}$ therefore becomes a random walk, and so the parameter a_0 may also be restricted to zero. The intuition for the latter restriction is that setting $\tilde{\kappa}_1 = 0$ converts $x_{1,t}$ into a pure Level state variable that represents the expected value, as at time t , of the long-run (infinite-maturity) short rate. Hence, a_0 is no longer required in the $\mathbb{E}_t[r(t + \tau)]$ component of the shadow-GATSM(2) forward rate equation.³⁰

Appendix G shows that the $\tilde{\kappa}_1 = 0$ restriction is readily imposed by applying the operator $\lim_{\tilde{\kappa}_1 \rightarrow 0}$ to the relevant components of equations 56 and 60. I have also imposed $a_0 = 0$ as noted, and I set $\tilde{\kappa}_2 = \phi$ for notational convenience. The final result is that the shadow-GATSM(2) forward rate expression becomes the two-factor arbitrage-free Nelson and Siegel (1987) model, or hereafter AFNSM(2), in its forward rate form, i.e.:

$$\begin{aligned}
 f(x_t, \mathbb{A}, u) &= x_{1,t} + x_{2,t} \cdot \exp(-\phi u) \\
 &\quad - \sigma_1^2 \cdot \tau^2 - \sigma_2^2 \cdot \frac{1}{2} [G(\phi, u)]^2 - \rho \sigma_1 \sigma_2 \cdot \tau \cdot G(\phi, u) \quad (61)
 \end{aligned}$$

³⁰Specifically, $\lim_{\tau \rightarrow \infty} \mathbb{E}_t[r(t + \tau)] = a_0$ for the shadow-GATSM(2) and $\lim_{\tau \rightarrow \infty} \mathbb{E}_t[r(t + \tau)] = x_1(t)$ for the restricted shadow-GATSM(2).

where 1 and $\exp(-\phi u)$ are the Nelson and Siegel (1987) Level and Slope forward rate components. (In the following section I introduce the third [Bow] Nelson and Siegel (1987) component.) The annualized option volatility becomes:

$$\omega(\mathbb{A}, u) = \sqrt{\sigma_1^2 \cdot u + \sigma_2^2 \cdot G(2\phi, u) + 2\rho\sigma_1\sigma_2 \cdot G(\phi, u)} \quad (62)$$

Using these results in conjunction with the state space representation already given in section 5.1 completes what I will hereafter call the CAB-AFNSM(2). I estimate the CAB-AFNSM(2) using the longest span of GSW30 data available, which is from November 1985 (the first observation in the GSW data set) to July 2013 (the last month-end data point available at the time of writing), and the GSW10 data over that same sample period.

Figure 4 plots the SSRs estimated using the GSW10 data set with the IEKF(1e-8) introduced earlier, the EKF, and the IEKF(1). The results clearly show that estimates obtained with the EKF and the IEKF(1) can be very misleading compared to those obtained with the IEKF(1e-8); in this case the SSR estimates remain only mildly negative while the IEKF(1e-8) SSR estimates adopt larger negative values. The IEKF(2) produced results closer to the IEKF(1e-8), but still with some material differences.³¹

Indeed, the IEKF(1e-8) required up to 10 iterations to obtain the given tolerance in periods with negative SSRs, compared to only several iterations in periods that produced positive shadow short rates. Estimations of the CAB-AFNSM(2) with the GSW30 data set showed similar divergences by estimation method to those presented in figure 4, and estimations of the CAB-GATSM(2) in section 5.2 and the restricted three-factor CAB-GATSM in section 5.4 also showed very material divergences by estimation method.

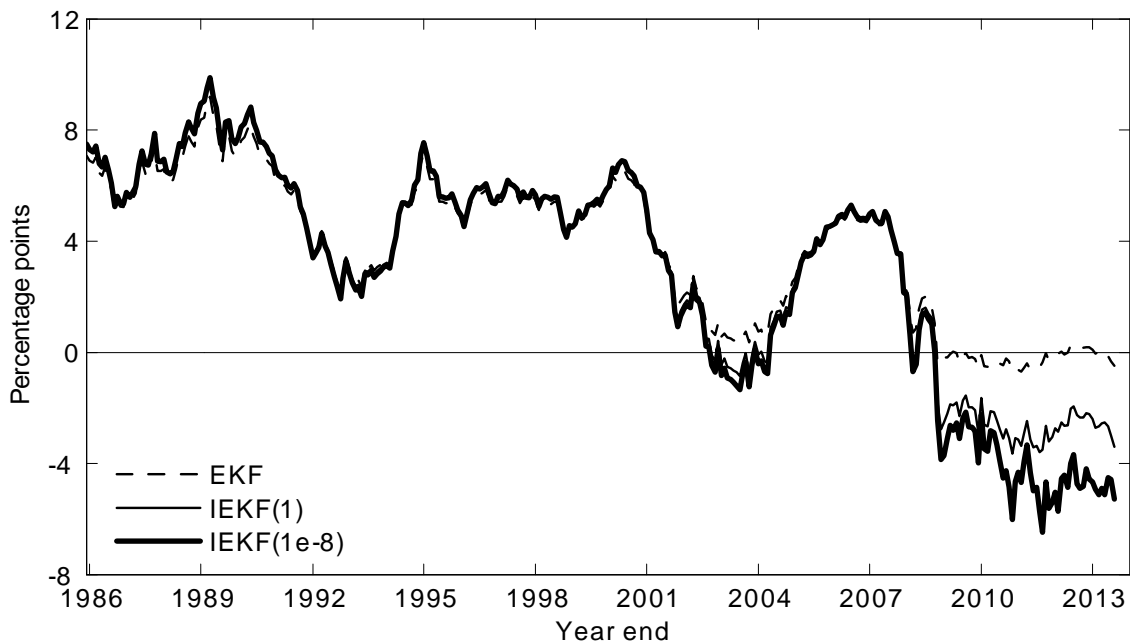


Figure 4: CAB-AFNSM(2) shadow short rates estimated from the GSW10 data set using the EKF, the IEKF with one additional iteration relative to the EKF, or the IEKF iterated until all elements of $|x_{t,i+1}^+ - x_{t,i+1}|$ are less than 1e-8.

³¹Statistically, the log-likelihood value for the IEKF(1e-8) estimate is 15052.3, while the EKF and IEKF(1) to IEKF(3) log-likelihood values are respectively lower by 257.3, 123.1, 20.2, and 2.0.

To provide some intuition for the variability of CAB-GATSM results by the method of estimation, figure 5 plots the SSR and the Level and Slope state variables for the GSW30 IEKF(1e-8) estimation (but the IEKF(1e-8) results from any model or data set could have been used). First, note that the Slope state variable varies by more than the Level state variable, and so makes the largest contribution to variation in the SSR.³² Second, because the non-linearity of the measurement equation will also be dominated by the Slope state variable, the non-linearity may effectively be viewed as a function of the estimated SSR. Third, the magnitudes of negative SSR estimates relate directly to how much the ZLB is constraining the yield curve data (i.e. a negative SSR essentially measures the difference between the ZLB [the lowest actual short rate in practice] and where the short rate would otherwise be in the absence of physical currency).

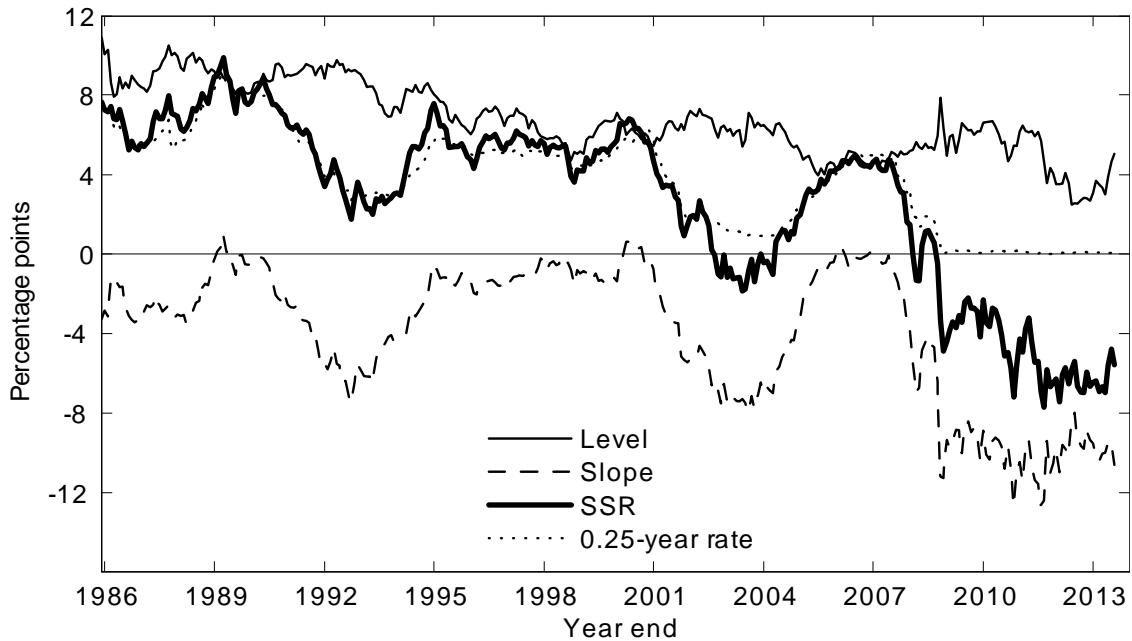


Figure 5: CAB-AFNSM(2) Level and Slope state variables, and shadow short rates estimated using the GSW30 data set and the IEKF(1e-8).

Bearing these ideas in mind, the non-linearity of the measurement equation should be small when the SSR is positive (i.e. conventional monetary policy periods, when all interest rates on the yield curve are materially positive), and more significant as the SSR becomes negative (i.e. unconventional monetary policy periods, when short- and mid-maturity interest rates are materially constrained by the ZLB). The EKF and lower-iteration IEKF estimations will only be suitable for mildly non-linear systems,³³ and so their performance will become relatively worse in unconventional monetary

³²The relatively low variability of the Level state variable accords with its interpretation as the long-horizon expectation of the SSR/actual short rate. As a related aside, from a macro-finance perspective, the long-horizon expected SSR/actual short rate should in principle be related to relatively slow-moving fundamental macroeconomic variables, such as long-run potential growth and expected inflation. Diebold, Rudebusch, and Aruoba (2006) offer empirical support for the latter relationship.

³³See, for example, Grewal and Andrews (2008) section 7.2 and Durbin and Koopman (2012) ch. 7 for further discussion on non-linear filtering in general, and James and Webber (2000) p. 527 for a comment on non-linear term structure models.

policy periods. Conversely, the IEKF(1e-8) estimations should be more robust to higher degrees of non-linearity.³⁴

5.4 Restricted CAB-GATSM(3)

In this section I develop a restricted three-factor CAB-GATSM, which uses the restricted GATSM(3) to represent the shadow term structure. The maximally identifiable GATSM(3) has 22 free parameters, and I follow Christensen, Diebold, and Rudebusch (2011) to impose the three additional restrictions that replicate the three-factor AFNSM, or AFNSM(3). Specifically, relative to maximally identifiable three-factor CAB-GATSM, I set $a_0 = 0$ and define the mean-reversion matrix $\tilde{\kappa}$ with a single parameter ϕ , i.e.:

$$\tilde{\kappa} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \phi & -\phi \\ 0 & 0 & \phi \end{bmatrix} \quad (63)$$

rather than $\tilde{\kappa} = \text{diag}[\tilde{\kappa}_1, \tilde{\kappa}_2, \tilde{\kappa}_3]$. Therefore, the eigenvalues of $\tilde{\kappa}$ are $\tilde{\kappa}_1 = 0$ and a repeated eigenvalue of ϕ .

The remaining free parameters are κ (nine parameters, because Γ contributes nine free parameters), θ (three parameters), and σ is defined with the six parameters σ_1 , σ_2 , σ_3 , ρ_{12} , ρ_{13} , and ρ_{23} , i.e.:

$$\sigma' = \begin{bmatrix} \sigma_1 & 0 & 0 \\ \sigma_2 \rho_{12} & \sigma_2 \sqrt{1 - \rho_{12}^2} & 0 \\ \sigma_3 \rho_{13} & \sigma_3 \frac{\rho_{23} - \rho_{12} \rho_{13}}{\sqrt{1 - \rho_{12}^2}} & \sigma_3 \sqrt{1 - \rho_{13}^2 - \frac{(\rho_{23} - \rho_{12} \rho_{13})^2}{1 - \rho_{12}^2}} \end{bmatrix} \quad (64)$$

Regarding the remaining parameters, I follow Christensen, Diebold, and Rudebusch (2011) in setting $\tilde{\theta} = 0$, and $b_0 = [1, 1, 0]'$. Appendix H details the derivation of the shadow-AFNSM(3) forward rate equation, with the final result:

$$\begin{aligned} f(x_t, \mathbb{A}, u) &= x_{1,t} + x_{2,t} \cdot \exp(-\phi u) - x_{3,t} \cdot \phi u \exp(-\phi u) \\ &\quad - \sigma_1^2 \cdot \frac{1}{2} u^2 - \sigma_2^2 \cdot \frac{1}{2} [G(\phi, u)]^2 - \rho \sigma_1 \sigma_2 \cdot u G(\phi, u) \\ &\quad - \sigma_3^2 \cdot \frac{1}{2} [F(\phi, u)]^2 + \rho_{13} \sigma_1 \sigma_3 \cdot u F(\phi, u) \\ &\quad + \rho_{23} \sigma_2 \sigma_3 \cdot G(\phi, u) F(\phi, u) \end{aligned} \quad (65)$$

where:

$$F(\phi, \tau) = G(\phi, \tau) - \tau \exp(-\phi \tau) \quad (66)$$

This is the AFNSM(3) in its forward rate form, where 1, $\exp(-\phi u)$, and $\phi \tau \exp(-\phi \tau)$ are respectively the Nelson and Siegel (1987) Level, Slope, Bow forward rate components.

³⁴It would also be useful to apply other non-linear methods in future work, such as the unscented Kalman filter and the particle filter, to assess if the IEKF subject to a minimum tolerance is generally robust enough for routine CAB-GATSM estimations.

The CAB-AFNSM(3) annualized option volatility is derived in appendix H as:

$$\begin{aligned}
[\omega(\mathbb{A}, u)]^2 &= \sigma_1^2 \cdot \tau + \sigma_2^2 \cdot G(2\phi, u) + 2\rho_{12}\sigma_1\sigma_2G(\phi, u) \\
&\quad + \sigma_3^2 \cdot \frac{1}{2} [F(2\phi, u) - \phi\tau^2 \exp(-2\phi u)] \\
&\quad - 2\rho_{13}\sigma_1\sigma_3F(\phi, u) - \rho_{23}\sigma_2\sigma_3F(2\phi, u)
\end{aligned} \tag{67}$$

Using these results in conjunction with the state space representation already given in section 5.1 completes the CAB-AFNSM(3) specification. Figure 6 plots the CAB-AFNSM(3) SSRs estimated from the GSW10 and GSW30 data sets with the IEFK(1e-8) method.

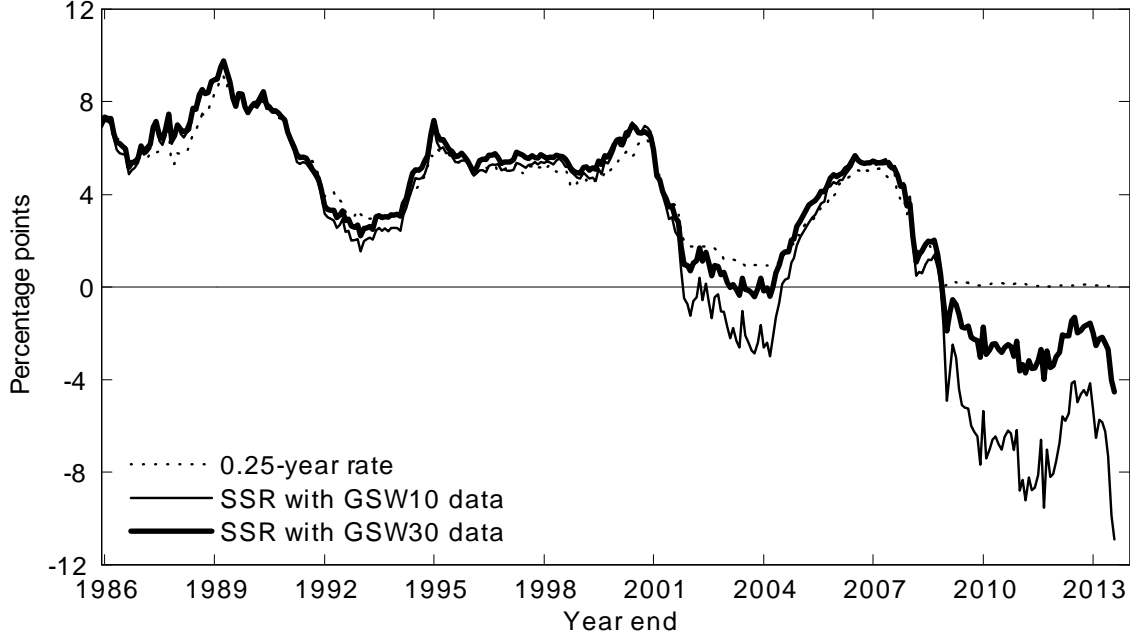


Figure 6: CAB-AFNSM(3) shadow short rates estimated with the GSW10 and GSW30 data sets and the IEFK(1e-8)

The two SSRs are obviously distinctly different, and the GSW10 estimates appear unusually negative compared to all other results presented in this article so far. The explanation for that result is that the CAB-AFNSM(3) appears overly flexible for the maturity span of 10 years. More precisely, the GAB-GATSM(3) allows an extra degree of flexibility for the expected path of the shadow short rate compared to the current SSR, i.e. the first line of equation 65 is:

$$\tilde{\mathbb{E}}_t[r(t + \tau)] = x_{1,t} + x_{2,t} \cdot \exp(-\phi\tau) - x_{3,t} \cdot \phi\tau \exp(-\phi\tau) \tag{68}$$

while the CAB-AFNSM(3) SSR at any point in time is $r_t = b'_0 x_t = x_{1,t} + x_{2,t}$. Therefore, the CAB-AFNSM(3) allows the current level of the SSR to differ from the short-to mid-horizon expectations of the SSR, with the difference dependent on the estimate of the Bow state variable $x_{3,t}$. For the GSW10 data set, $x_{3,t}$ reaches almost 20 percent at times after late-2008, and the Slope state variable $x_{2,t}$ reaches almost -15 percent at corresponding times. These are large offsetting magnitudes suggestive of overfitting. Conversely, the CAB-AFNSM(3) GSW30 Slope and Bow estimates always retain magnitudes below 10 percent, and so using data with longer maturities appears to mitigate the potential for overfitting within the CAB-AFNSM(3).

5.5 Discussion of results

Three broad results that are apparent from sections 5.2 to 5.4 are that the estimated SSR results are sensitive to: (1) the data set used for the estimation; (2) the model specification; and (3) the method of estimation.³⁵ Aspect (3) may be resolved in principle by using an estimation method that appropriately allows for higher degrees of non-linearity in the measurement equation, which occurs when the data becomes more constrained by the ZLB. However, aspects (1) and (2) are intrinsic to choices by the researcher, thus highlighting that any results and comparisons of estimated CAB-GATSM models need to be caveated explicitly in the context of those choices. The wider question of which data set and model specification is most appropriate, in general or for a given application, remains subject to further investigation, although I make a few brief related comments below.

The three observations regarding applications of CAB-GATSMs also hold implications for the application of Black models. In particular, estimates of Black models based on EKF or linearized versions of the Kalman filter will likely be subject to the same estimation issues illustrated here. Again, the precision and convergence of state variable estimates is something that could be resolved with appropriate estimation techniques. However, because Black framework implementations use numerical methods, the larger number of numerical evaluations required for implementing and assessing alternative estimation techniques may prove practically infeasible.

The CAB-GATSM results also offer some evidence on two important issues that, due to space constraints, I cannot address fully in this article, i.e.: (1) how closely the CAB-GATSM results approximate those that would otherwise have been obtained using the Black framework; and (2) whether the SSR estimates for any model may fulfill the interpretation of an indicator of the stance of monetary policy across conventional and unconventional monetary policy environments, as suggested in Krippner (2012 and 2013a).

On the first aspect, the close correspondence between the Black-GATSM(2) and the CAB-GATSM(2) SSRs in figure 2 suggests that the CAB-GATSM approximation appears practically adequate in that particular case. Therefore, it is not unreasonable to anticipate that a similar degree of approximation may carry through to the other estimated models. Further investigation will be required, of course, but other reported results to date are encouraging.³⁶

³⁵In other results not presented in this article due to space considerations, I found some sensitivity with respect to the sample period and to the way the data sets were generated (for example, when using Bloomberg zero coupon data which is generated by bootstrapping rather than the GSW data which is generated by curve fitting). I also found very different results when using the EKF or low-iteration IEKFs to estimate a given CAB-GATSM specification with the same data but at different frequencies (e.g. monthly, weekly, and daily). That variance of results was the initial reason I opted for an IEKF with a tight precision. Using the IEKF(1e-8), the monthly results results in the present article are very close to the results obtained with different data frequencies.

³⁶For example, Christensen and Rudebusch (2013) p. 22 presents CAB-AFNSM(2) SSRs for Japan (albeit estimated using the EKF) that are similar to the Kim and Singleton (2012) Black-GATSM(2) estimates, and also see the discussion at the end of section 3.1 in the present article. As a related aside, Christensen and Rudebusch (2013) p. 7 makes the very pertinent point: “Although many shadow-rate term structure papers start with a theoretically consistent model, various simplifications are made to facilitate empirical implementation.” Regarding the Black applications in the present article, Kim and Singleton (2012) p. 46 notes the potential for discretization errors from the pragmatic choices for their

On the second aspect, the CAB-AFNSM(2) GSW30 SSR estimates have a profile and magnitude similar to the illustrative results reported in Krippner (2012),³⁷ and are therefore broadly consistent with the timing of unconventional monetary policy events noted in that article. Specifically, the SSR first went negative around the announcement of the first large scale asset purchase (LSAP) program in November 2008, became more negative after the second LSAP program was pre-announced in August 2010, and adopted the most negative values over 2011 to 2013 as forward guidance statements were issued. In addition, the SSR estimate has risen since May 2013, following Federal Reserve comments about tapering the third LSAP program that was originally introduced in September and October 2012. In general, the SSR results based on GSW30 data are more stable and consistent across different CAB-GATSM specifications than the GSW10 results. Further investigation of whether and how well SSRs from particular CAB-GATSM specifications and data sets match unconventional monetary events would be required before any conclusions can be drawn, but it seems premature to adopt the blanket recommendation from Christensen and Rudebusch (2013) p. 23-24 against using estimated SSRs as a monetary policy indicator. Conversely, the SSR estimates from the CAB-GATSM specifications using the GSW10 data sets already appear either inconsistent with unconventional monetary events and/or quite variable, particularly the CAB-AFNSM(3) results, and so they appear to already rule themselves out as reliable monetary policy indicators.

6 Conclusion

In this article I have developed a tractable generic framework for imposing a zero lower bound (ZLB) on Gaussian affine term structure models (GATSMs). My “currency-adjusted bond” or CAB-GATSM framework is an approximation to the Black framework in the sense that it uses shadow short rates to discount future contingent cash-flows, while the Black framework uses own-model ZLB rates for discounting. However, the approximation appears to be acceptable in practice; i.e. encouragingly, the two-factor CAB-GATSM shadow short rates I estimate from the data described in Ichiue and Ueno (2013) are very similar to the two-factor Black-GATSM results from that same paper

More generally, the CAB-GATSM framework is much more tractable expression than the Black framework, because it produces closed form forward rate expressions for any specification. As I have shown in the empirical application of this article, that tractability readily allows the estimation of models with more than two factors, on data sets with long maturities, and using robust methods that potentially requires

finite difference application, Ichiue and Ueno (2013) p. 11 notes that the Jensen effect is ignored in their lattice application, and Bauer and Rudebusch (2013) uses pre-estimated parameters for their Monte Carlo simulation application. While all offer evidence that the effect of the simplifications should be small, they are nevertheless approximations to the idealized Black framework, and their effect should ideally be assessed as a matter of course.

³⁷The results in Krippner (2012) obtained from model with calibrated parameters and non-linear least squares estimation of the state variables from each month-end observation of yield curve data. Krippner (2012) p. 25 reports a mean of -4.97 percent for the SSR from August 2010 to the end of the sample (January 2012), while the mean of the GSW30 SSR results in the present article for the same period is -5.68 percent.

many iteration to achieve convergence for the state variable estimates.

References

- Babbs, S. and K. Nowman (1998). An application of generalized Vasicek term structure models to the UK gilt-edged market: a Kalman filtering analysis. *Applied Financial Economics* 8, 637–644.
- Bauer, M. and G. Rudebusch (2013). Monetary policy expectations at the zero lower bound. *Working Paper, Federal Reserve Bank of San Francisco*.
- 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*.
- Brace, A., D. Gatarek, and M. Musiela (1997). The market model of interest rate dynamics. *Mathematical Finance* 7, 127–155.
- Bullard, J. (2012). Shadow Interest Rates and the Stance of U.S. Monetary Policy. *Speech at the Annual Conference, Olin Business School, Washington University in St. Louis, 8 November 2012*.
- Bullard, J. (2013). Perspectives on the Current Stance of Monetary Policy. *Speech at the NYU Stern, Center for Global Economy and Business, 21 February 2013*.
- Chen, R. (1995). A two-factor, preference-free model for interest rate sensitive claims. *Journal of Futures Markets* 15(3), 345–372.
- Chen, R. and L. Scott (1992). Pricing interest rate options in a two-factor Cox–Ingersoll–Ross model of the term structure. *Review of Financial Studies* 5(4), 613–636.
- Cheridito, P., D. Filipović, and R. Kimmel (2007). Market price of risk specifications for affine models: theory and evidence. *Journal of Financial Economics* 83, 123–170.
- 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.
- Christensen, J. and G. Rudebusch (2013). Estimating shadow-rate term structure models with near-zero yields. *Working Paper, Federal Reserve Bank of San Francisco*.
- Cox, J., J. Ingersoll, and S. Ross (1985). 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 G. Rudebusch (2013). *Yield Curve Modeling and Forecasting: The Dynamic Nelson–Siegel Approach*. Princeton University Press.

- Diebold, F., G. Rudebusch, and S. Aruoba (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics* 131, 309–338.
- Duffee, G. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance* 57(1), 405–443.
- Durbin, J. and S. Koopman (2012). *Time Series Analysis by State Space Methods, Second Edition*. Oxford University Press.
- 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.
- Grewal, M. and A. Andrews (2008). *Kalman Filtering, Third Edition*. Wiley and Sons.
- 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 (2012). Identification and estimation of affine-term-structure models. *Journal of Econometrics* 168(2), 315–331.
- Heath, D., R. Jarrow, and A. Morton (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* 60(1), 77–106.
- Higam, N. (2008). *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics.
- 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.
- Ichiue, H. and Y. Ueno (2013). Estimating term premia at the zero lower bound: an analysis of Japanese, US, and UK yields. *Working Paper, Bank of Japan* 13-E-8.
- 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 (2012). Term structure models and the zero bound: an empirical investigation of Japanese yields. *Journal of Econometrics* 170(1), 32–49.
- Klebaner, F. (2005). *Introduction to Stochastic Calculus With Applications, Second Edition*. Imperial College Press.
- Krippner, L. (2012). Modifying Gaussian term structure models when interest rates are near the zero lower bound. *Discussion Paper, Reserve Bank of New Zealand* DP2012/02.
- Krippner, L. (2013). Measuring the stance of monetary policy in zero lower bound environments. *Economics Letters* 118, 135–138.

- Meucci, A. (2010). Review of statistical arbitrage, cointegration, and multivariate Ornstein-Uhlenbeck. *Working Paper, SYMMYS*.
- Moler, C. and C. Van Loan (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review* 45(1), 3–48.
- Nelson, C. and A. Siegel (1987). Parsimonious modelling of yield curves. *Journal of Business* October, 473–489.
- 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.
- Richard, S. (2013). A non-linear macroeconomic term structure model. *Working Paper, University of Pennsylvania*.
- Rudebusch, G. (2010). Macro-finance models of interest rates and the economy. *The Manchester School Supplement* 2010, 25–52.
- Simon, D. (2006). *Optimal State Estimation*. Wiley and Sons.
- Singleton, K. (2006). *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. Princeton University Press.
- Tchuindjo, L. (2008). Factor’s correlation in the Heath-Jarrow-Morton interest rate model. *Applied Stochastic Models in Business and Industry* 24, 359–368.
- Tchuindjo, L. (2009). An extended Heath-Jarrow-Morton risk-neutral drift. *Applied Financial Mathematics Letters* 22, 396–400.
- 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.
- Wu, C. and F. Xia (2013). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Working Paper*.

A General CAB forward rate and short rate

Appendix A.1 contains the detailed derivation of the general CAB forward rate expression $\underline{f}(t, \tau)$ in equation 25. Appendix A.2 contains the detailed results for the general CAB short rate $\underline{r}(t)$ as summarized in equation 30. I note the use of elementary calculus and limit operations for the benefit of the reader, but without reference.

A.1 General CAB forward rate derivation

$$\begin{aligned}
\underline{f}(t, \tau) &= \lim_{\delta \rightarrow 0} \left\{ -\frac{1}{\delta} \log \left[\frac{P(t, \tau + \delta) - C(t, \tau, \delta)}{P(t, \tau)} \right] \right\} \\
\langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta) - C(t, \tau, \delta)}{P(t, \tau)} \right] \right\}}{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \delta \right\}} \\
&= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log [P(t, \tau + \delta) - C(t, \tau, \delta)] - \frac{d}{d\delta} \log [P(t, \tau)] \right\}}{\lim_{\delta \rightarrow 0} \{1\}} \\
&= -\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log [P(t, \tau + \delta) - C(t, \tau, \delta)] \right\} \\
\langle \text{Chain rule} \rangle &= -\lim_{\delta \rightarrow 0} \left\{ \frac{d}{dx} \log [x] \cdot \frac{d}{d\delta} [P(t, \tau + \delta) - C(t, \tau, \delta)] \right\} \\
&\quad \left\langle x = P(t, \tau + \delta) - C(t, \tau, \delta); \frac{d}{dx} \log [x] = \frac{1}{x} \right\rangle \\
&= -\lim_{\delta \rightarrow 0} \left\{ \frac{1}{P(t, \tau + \delta) - C(t, \tau, \delta)} \cdot \frac{d}{d\delta} [P(t, \tau + \delta) - C(t, \tau, \delta)] \right\} \\
\langle \text{Limit rules} \rangle &= -\lim_{\delta \rightarrow 0} \left\{ \frac{1}{P(t, \tau + \delta) - C(t, \tau, \delta)} \right\} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} [P(t, \tau + \delta)] \right\} \\
&\quad + \lim_{\delta \rightarrow 0} \left\{ \frac{1}{P(t, \tau + \delta) - C(t, \tau, \delta)} \right\} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} [C(t, \tau, \delta)] \right\} \quad (69)
\end{aligned}$$

where ‘‘Limit rules’’ refers to the standard rules of calculus for manipulating limits of analytic functions, in this case the product limit rule, i.e. $\lim_{x \rightarrow a} \{f(x) \cdot g(x)\} = \lim_{x \rightarrow a} f(x) \cdot \lim_{x \rightarrow a} g(x)$ and the sum/difference limit rule, i.e. $\lim_{x \rightarrow a} \{f(x) \pm g(x)\} = \lim_{x \rightarrow a} f(x) \pm \lim_{x \rightarrow a} g(x)$. In what immediately follows, I also use the quotient limit rule, i.e. $\lim_{x \rightarrow a} \{f(x)/g(x)\} = \lim_{x \rightarrow a} f(x) / \lim_{x \rightarrow a} g(x)$ if $\lim_{x \rightarrow a} g(x) \neq 0$, hence:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{1}{P(t, \tau + \delta) - C(t, \tau, \delta)} \right\} &= \lim_{\delta \rightarrow 0} \{ [P(t, \tau + \delta) - C(t, \tau, \delta)]^{-1} \} \\
\langle \text{Limit rules} \rangle &= \left[\lim_{\delta \rightarrow 0} \{ P(t, \tau + \delta) - C(t, \tau, \delta) \} \right]^{-1} \\
\langle \text{Limit rules} \rangle &= \left[\lim_{\delta \rightarrow 0} \{ P(t, \tau + \delta) \} - \lim_{\delta \rightarrow 0} \{ C(t, \tau, \delta) \} \right]^{-1} \\
&= [P(t, \tau) - 0]^{-1} \\
&= \frac{1}{P(t, \tau)} \quad (70)
\end{aligned}$$

where the result $\lim_{\delta \rightarrow 0} \{C(t, \tau, \delta)\} = 0$ 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:

$$C(t, \tau, \delta) = \tilde{\mathbb{E}}_t \left\{ \exp \left(- \int_0^\tau r(t+u) du \right) \cdot \max [P(t+\tau, \delta) - 1, 0] \right\} \quad (71)$$

and the result of zero follows because the limit of the payoff component in all states is zero, i.e.:

$$\begin{aligned} \lim_{\delta \rightarrow 0} \{ \max [P(t+\tau, \delta) - 1, 0] \} &= \max \left[\lim_{\delta \rightarrow 0} \{P(t+\tau, \delta) - 1\}, 0 \right] \\ &= \max [P(t+\tau, 0) - 1, 0] \\ &= \max [1 - 1, 0] \\ &= 0 \end{aligned} \quad (72)$$

The term $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} P(t, \tau + \delta) \right\}$ may be evaluated using the definition of differentiation as a limit with respect to a dummy variable ξ , reversing the order of the two limit calculations, and then expressing the result as a differential again, i.e.:

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

Therefore, equation 69 becomes:

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

where $f(t, \tau)$ is obtained from the standard term structure relationship between forward rates and bond prices, i.e.:

$$\begin{aligned} f(t, \tau) &= -\frac{d}{d\tau} \log [P(t, \tau)] \\ \langle \text{Chain rule} \rangle &= -\frac{d}{dx} \log [x] \cdot \frac{d}{d\tau} P(t, \tau) \\ &\quad \left\langle x = P(t, \tau); \frac{d}{dx} \log [x] = \frac{1}{x} \right\rangle \\ &= -\frac{1}{P(t, \tau)} \cdot \frac{d}{d\tau} P(t, \tau) \end{aligned} \quad (74)$$

and I leave $z(t, \tau)$ implicit at this stage, i.e.:

$$z(t, \tau) = \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \delta)}{P(t, \tau)} \right] \right\} \quad (75)$$

pending the GATSM specification for $C(t, \tau, \delta)$ and $P(t, \tau)$ in section 4.

A.2 General CAB short rate derivation

I begin from the following standard term structure relationship:

$$\begin{aligned}
\underline{r}(t) &= \lim_{\tau \rightarrow 0} \{f(t, \tau)\} \\
&= \lim_{\tau \rightarrow 0} \{f(t, \tau) + z(t, \tau)\} \\
\langle \text{Limit rules} \rangle &= \lim_{\tau \rightarrow 0} \{f(t, \tau)\} + \lim_{\tau \rightarrow 0} \left\{ \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \delta)}{P(t, \tau)} \right] \right\} \right\} \\
&= f(t, 0) + \lim_{\tau \rightarrow 0} \left\{ \lim_{\delta \rightarrow 0} \left\{ \lim_{\xi \rightarrow 0} \left[\frac{C(t, \tau, \delta + \xi) - C(t, \tau, \delta)}{P(t, \tau) \cdot \xi} \right] \right\} \right\} \\
&= r(t) + \lim_{\delta \rightarrow 0} \left\{ \lim_{\xi \rightarrow 0} \left\{ \lim_{\tau \rightarrow 0} \left[\frac{C(t, \tau, \delta + \xi) - C(t, \tau, \delta)}{P(t, \tau) \cdot \xi} \right] \right\} \right\} \\
&= r(t) + \lim_{\delta \rightarrow 0} \left\{ \lim_{\xi \rightarrow 0} \left[\frac{C(t, 0, \delta + \xi) - C(t, 0, \delta)}{1 \cdot \xi} \right] \right\} \\
&= r(t) + \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} C(t, 0, \delta) \right\} \tag{76}
\end{aligned}$$

where the results $\lim_{\tau \rightarrow 0} \{f(t, \tau)\} = f(t, 0) = r(t)$ are term structure definitions, and I have used the definition of differentiation as a limit in the intermediate steps.

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

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} C(t, 0, \delta) \right\} &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \max \{P(t, \delta) - 1, 0\} \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \max \left[\frac{d}{d\delta} [P(t, \delta) - 1], 0 \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \max \left[\frac{d}{d\delta} P(t, \delta), 0 \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \{ \max [-P(t, \delta) \cdot f(t, \delta), 0] \} \\
\langle \text{Limit rules} \rangle &= \max \left[\lim_{\delta \rightarrow 0} \{-P(t, \delta) \cdot f(t, \delta)\}, 0 \right] \tag{77}
\end{aligned}$$

where the result $\frac{d}{d\delta} P(t, \delta) = -P(t, \delta) \cdot f(t, \delta)$ is apparent from re-arranging equation 74 and substituting δ for τ , and the final step uses the limit rule for functions, i.e. $\lim_{x \rightarrow a} \{g[f(x)]\} = g[\lim_{x \rightarrow a} f(x)]$. Finally, applying the product limit rule gives:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \{-P(t, \delta) \cdot f(t, \delta)\} &= -\lim_{\delta \rightarrow 0} \{P(t, \delta)\} \cdot \lim_{\delta \rightarrow 0} \{f(t, \delta)\} \\
&= -1 \cdot f(t, 0) \\
&= -r(t) \tag{78}
\end{aligned}$$

Substituting the option effect result back through to the original expression in equation 76 gives the final result:

$$\begin{aligned}
\underline{r}(t) &= r(t) + \max \{-r(t), 0\} \\
&= \max \{0, r(t)\} \tag{79}
\end{aligned}$$

B Shadow-GATSM forward rate derivation

Appendix B.1 derives the shadow-GATSM bond price expression $P(t, \tau)$ given in section 3.2, and appendix B.2 uses that derivation to obtain the shadow-GATSM forward rate expression $f(t, \tau)$ also given in section 3.2. Appendix B.3 derives the option effect expression $z(t, \tau)$ in section 3.3, and appendix B.3 derives the annualized option volatility expression $\omega(\tau)$.

B.1 Shadow-GATSM bond prices

Dai and Singleton (2002) p. 438 gives the following expression for GATSM bond prices:

$$P(t, \tau) = \exp[-A(\tau) - B(\tau)'x(t)] \quad (80)$$

where:

$$B(\tau) = [I - \exp(-\tilde{\kappa}'\tau)] \tilde{\kappa}'^{-1}b_0 \quad (81)$$

with I the $N \times N$ identity matrix. Dai and Singleton (2002) also provide an expression for $A(\tau)$ that applies to GATSMs with distinct eigenvalues for $\tilde{\kappa}$, such as the two-factor shadow-GATSMs that I specify in sections 5.2 and 5.3.³⁸ I use a more general form for $A(\tau)$ that accommodates cases where $\tilde{\kappa}$ does not have distinct eigenvalues, such as the restricted three-factor CAB-GATSM that I develop in section 5.4. Hence:

$$A(\tau) = \frac{1}{\tau} \int_0^\tau \frac{d}{dv} A(v) dv \quad (82)$$

where:

$$\frac{d}{dv} A(v) = a_0 + \int_0^v [\mathfrak{h}(b_0, \tilde{\kappa}, v-s)]' \sigma' \left[\int_s^v \sigma [\mathfrak{h}(b_0, \tilde{\kappa}, u-s)] du \right] ds \quad (83)$$

and $[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]$ the vector function of forward rate loadings:

$$[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)] = b_0' \exp(-\tilde{\kappa}\tau) \quad (84)$$

My expression for $A(\tau)$ arises from the Heath, Jarrow, and Morton (1992) framework with an explicit allowance for correlations as introduced by Tchuindjo (2008, 2009), i.e.:

$$\frac{d}{d\tau} A(\tau) = a_0 + \int_0^\tau \left\langle V(\tau, s), \rho \int_s^\tau V(u, s) du \right\rangle ds \quad (85)$$

where $V(\cdot, \cdot)$ is the vector function of forward rate volatilities, ρ is the $N \times N$ matrix of the correlations between the factor innovations, and $\langle \cdot, \cdot \rangle$ is the vector inner product. In the GATSM context, $V(\tau, s) = [\sigma_1 \cdot b_{0,1} \cdot \mathfrak{h}_1(\tilde{\kappa}, \tau-s), \dots, \sigma_N \cdot b_{0,N} \cdot \mathfrak{h}_N(\tilde{\kappa}, \tau-s)]'$. I express the vector inner product explicitly as $\langle a, b \rangle = a'b$, and equivalently re-express the standard deviations and correlation matrix in the form $\sigma'\sigma$; i.e. $\sigma'\sigma = v'\rho v$, where $v = [\sigma_1, \dots, \sigma_N]'$. Therefore, equation 85 can be re-expressed in terms of forward

³⁸Indeed, the subsequent results I obtain for those models are identical to those that would otherwise have been obtained using the Dai and Singleton (2002) expression for bond prices, and my results are readily extendable to replicate the Dai and Singleton (2002) results for N factors (as I note at the end of appendix C.2).

rate volatility loadings $\sigma(\tau)$ that already embed the correlations/covariances via the constant volatility matrix σ , i.e.:

$$\sigma(\tau) = \sigma[\mathbb{1}(b_0, \tilde{\kappa}, \tau)] \quad (86)$$

which is the notation I have used in equation 83.

B.2 Shadow-GATSM forward rates

Shadow-GATSM forward rates may be derived using the standard relationship between forward rates and bond prices, i.e.:

$$\begin{aligned} f(t, \tau) &= -\frac{d}{d\tau} \log P(t, \tau) \\ &= \frac{d}{d\tau} [A(\tau) + B(\tau)' x(t)] \\ &= \frac{d}{d\tau} A(\tau) + \frac{d}{d\tau} B(\tau)' x(t) \end{aligned} \quad (87)$$

The derivative for $A(\tau)$ with respect to time to maturity τ has already been given in equation 83 given how I have chosen to define $A(\tau)$. The derivative for $B(\tau)'$ with respect to time to maturity τ is:

$$\begin{aligned} \frac{d}{d\tau} B(\tau)' &= \frac{d}{d\tau} \{ [I - \exp(-\tilde{\kappa}'\tau)] \tilde{\kappa}'^{-1} b_0 \}' \\ &= \frac{d}{d\tau} \{ b_0' [\tilde{\kappa}'^{-1}]' [I - \exp(-\tilde{\kappa}'\tau)]' \}' \\ &= \frac{d}{d\tau} \{ b_0' \tilde{\kappa}^{-1} [I - \exp(-\tilde{\kappa}\tau)] \}' \\ &= -b_0' \tilde{\kappa}^{-1} \frac{d}{d\tau} \exp(-\tilde{\kappa}\tau) \\ &= -b_0' \tilde{\kappa}^{-1} \cdot -\tilde{\kappa} \exp(-\tilde{\kappa}\tau) \\ &= b_0' \exp(-\tilde{\kappa}\tau) \\ &= [\mathbb{1}(b_0, \tilde{\kappa}, \tau)] \end{aligned} \quad (88)$$

Substituting the expressions for $\frac{d}{d\tau} A(\tau)$ and $\frac{d}{d\tau} B(\tau)'$ into equation 87 therefore produces the final result as given in section 4.2, i.e.:

$$f(t, \tau) = a_0 + \int_0^\tau [\mathbb{1}(b_0, \tilde{\kappa}, \tau - s)]' \sigma' \left[\int_s^\tau \sigma[\mathbb{1}(b_0, \tilde{\kappa}, u - s)] du \right] ds + [\mathbb{1}(b_0, \tilde{\kappa}, \tau)]' x(t) \quad (89)$$

C CAB-GATSM option effect: limit derivation

This appendix provides a derivation of the option effect in section 4.3 that uses GATSM bond prices from Dai and Singleton (2002) p. 438, the generic closed-form analytic

expression for the price of call options on GATSM bonds from Filipović (2009) p. 109,³⁹ and my option effect expression in equation 26.

The expression for GATSM bond prices is already available in appendix B.1, and Filipović (2009) p. 109 gives the price of a call option with a strike price of 1 on a zero-coupon bond as:

$$C(t, \tau, \delta) = P(t, \tau + \delta) \Phi[d_1(t, \tau, \delta)] - P(t, \tau) \Phi[d_2(t, \tau, \delta)] \quad (90)$$

where $\Phi[\cdot]$ is the cumulative normal distribution function, and:

$$d_{1,2}(t, \tau, \delta) = \frac{1}{\Sigma(\tau, \delta)} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \pm \frac{1}{2} \Sigma(\tau, \delta) \quad (91)$$

with:

$$\Sigma(\tau, \delta) = \sqrt{\int_0^\tau \|\sigma(s, \delta)\|^2 ds} \quad (92)$$

and:

$$\sigma(s, \delta) = \int_s^{s+\delta} \sigma(u) du \quad (93)$$

Substituting the generic call option expression into equation 26 then gives the following result for the option effect:

$$z(t, \tau) = -f(t, \tau) \cdot \left(1 - \Phi \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \quad (94)$$

where $\Phi[\cdot]$ is the standard univariate cumulative normal distribution function and $\omega(\tau)$ is the annualized instantaneous option volatility derived in appendix B.4 as:

$$\begin{aligned} \omega(\tau) &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\delta} \Sigma(\tau, \delta) \right\} \\ &= \sqrt{\int_0^\tau [\mathfrak{h}(b_0, \tilde{\kappa}, s)]' \sigma' \sigma [\mathfrak{h}(b_0, \tilde{\kappa}, s)] ds} \end{aligned} \quad (95)$$

³⁹Filipović (2009) already assumes $t = 0$ (see Filipović (2009) footnote 1, p. 109) without loss of generality, so no change of variables is required to convert from time-of-maturity notation to my time-to-maturity notation. That is, T and S from Filipović (2009) are respectively equivalent to my τ and $\tau + \delta$. Also note that Chen (1995) provides the bond option expression for a two-factor GATSM, and notes that the results readily extend to N factors, but that is only the case if $\tilde{\kappa}$ has distinct eigenvalues. The generic Filipović (2009) expression accommodates GATSMs with repeated eigenvalues for $\tilde{\kappa}$, which accommodates the model in section 5.4. Christensen and Rudebusch (2013) explicitly derive the bond option expression for the same model, but there is no need to do so given the availability of the generic textbook result. Furthermore, as I establish in this appendix, and even more directly in section 4.3, only the volatility and forward rate expressions that underlie the bond option expression are required for the CAB-GATSM framework. Hence, there is no need to explicitly calculate the shadow-GATSM bond option expressions, as suggested by Christensen and Rudebusch (2013) p. 9, to evaluate the option effect expression for any CAB-GATSM; one should just use the GATSM volatility and forward rate expressions directly.

Equation 26 may then be evaluated as:

$$\begin{aligned}
z(t, \tau) &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{C(t, \tau, \delta)}{P(t, \tau)} \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta) \Phi[d_1(\cdot)] - P(t, \tau) \Phi[d_2(\cdot)]}{P(t, \tau)} \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \Phi[d_1(\cdot)] - \Phi[d_2(\cdot)] \right] \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \cdot \Phi[d_1(\cdot)] \right. \\
&\quad \left. + \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} \Phi[d_1(\cdot)] - \frac{d}{d\delta} \Phi[d_2(\cdot)] \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} \cdot \lim_{\delta \rightarrow 0} \{ \Phi[d_1(t, \tau, \delta)] \} \\
&\quad + \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} \Phi[d_1(\cdot)] - \frac{d}{d\delta} \Phi[d_2(t, \tau, \delta)] \right\} \quad (96)
\end{aligned}$$

The details for deriving the three limits required for equation 96 are contained in the following subsections, with the following respective results:

$$\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} = -f(t, \tau) \quad (97)$$

$$\lim_{\delta \rightarrow 0} \{ \Phi[d_1(t, \tau, \delta)] \} = 1 - \Phi \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \quad (98)$$

where $\omega(\tau)$ is annualized option volatility (which I derive in section B.4), and:

$$\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta} \Phi[d_1(\cdot)] - \frac{d}{d\delta} \Phi[d_2(\cdot)] \right\} = \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \quad (99)$$

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

$$z(t, \tau) = -f(t, \tau) \cdot \left(1 - \Phi \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)} \right]^2 \right) \quad (100)$$

C.0.1 $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\}$

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\}$ in equation 96 may be evaluated by first calculating the following derivative:

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

The expression $\frac{d}{d\delta}P(t, \tau + \delta)$ may be obtained using the result in equation 74 and some re-arrangement, i.e.:

$$\begin{aligned} f(t, \tau + \delta) &= -\frac{1}{P(t, \tau + \delta)} \frac{d}{d[\tau + \delta]} P(t, \tau + \delta) \\ \langle \text{Chain rule} \rangle &= -\frac{1}{P(t, \tau + \delta)} \frac{d}{d\delta} P(t, \tau + \delta) \frac{d\delta}{d[\tau + \delta]} \\ -f(t, \tau + \delta) \cdot P(t, \tau + \delta) \cdot \frac{d[\tau + \delta]}{d\delta} &= \frac{d}{d\delta} P(t, \tau + \delta) \end{aligned} \quad (101)$$

The final result is therefore:

$$\begin{aligned} \frac{1}{P(t, \tau)} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} P(t, \tau + \delta) \right\} &= \frac{1}{P(t, \tau)} \cdot \lim_{\delta \rightarrow 0} \left\{ -f(t, \tau + \delta) \cdot P(t, \tau + \delta) \cdot \frac{d[\tau + \delta]}{d\delta} \right\} \\ &= -\frac{1}{P(t, \tau)} \cdot \lim_{\delta \rightarrow 0} \{f(t, \tau + \delta)\} \cdot \lim_{\delta \rightarrow 0} \{P(t, \tau + \delta) \cdot 1\} \\ &= -f(t, \tau) \end{aligned} \quad (102)$$

C.0.2 $\lim_{\delta \rightarrow 0} \{\Phi[d_1(t, \tau, \tau + \delta)]\}$

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

$$\Phi[x] = \frac{1}{2} + \frac{1}{2} \text{erf} \left[\frac{1}{\sqrt{2}} x \right] \quad (103)$$

and the definition of the error function is itself:

$$\text{erf}[y] = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n [y]^{2n+1}}{n! (2n+1)} \quad (104)$$

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

$$\Phi[d_1(t, \tau, \delta)] = \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} d_1(t, \tau, \delta) \right]^{2n+1}}{n! (2n+1)} \quad (105)$$

and taking the limit gives:

$$\begin{aligned} \lim_{\delta \rightarrow 0} \Phi[d_1(\cdot)] &= \frac{1}{2} + \frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} d_1(\cdot) \right]^{2n+1}}{n! (2n+1)} \right\} \\ \langle \text{Limit rules} \rangle &= \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \cdot \lim_{\delta \rightarrow 0} \left\{ \left[\frac{1}{\sqrt{2}} d_1(\cdot) \right]^{2n+1} \right\}}{n! (2n+1)} \\ \langle \text{Limit rules} \rangle &= \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} \cdot \lim_{\delta \rightarrow 0} d_1(\cdot) \right]^{2n+1}}{n! (2n+1)} \end{aligned} \quad (106)$$

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

The expression for $\lim_{\delta \rightarrow 0} d_1(t, \tau, \delta)$ is derived directly from the expression in equation 91, i.e.:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} d_1(t, \tau, \delta) &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\Sigma(\tau, \delta)} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] + \frac{1}{2} \Sigma(\tau, \delta) \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\Sigma(\tau, \delta)} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} + \frac{1}{2} \lim_{\delta \rightarrow 0} \Sigma(\tau, \delta) \\
\langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\}}{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \delta) \right\}} + \frac{1}{2} \lim_{\delta \rightarrow 0} \Sigma(\tau, \delta) \quad (107)
\end{aligned}$$

where L'Hopital's rule has been applied to the first expression on the second line of equation 107 because $\lim_{\delta \rightarrow 0} \Sigma(\tau, \delta) = 0$ (which follows because the upper and lower limits of the integral of equation 93 would be identical) meaning the limit would otherwise be undefined (i.e. $\log [P(t, \tau + \delta) / P(t, \tau)] = \log [1] = 0$ and $\Sigma(\tau, 0) = 0$).

The expression $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\}$ required for equation 107 is:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log \left[\frac{P(t, \tau + \delta)}{P(t, \tau)} \right] \right\} &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} (\log [P(t, \tau + \delta)] - \log [P(t, \tau)]) \right\} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \log [P(t, \tau + \delta)] - \frac{d}{d\delta} \log [P(t, \tau)] \right\} \\
\langle \text{Chain rule} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d[\tau + \delta]} \log [P(t, \tau + \delta)] \cdot \frac{d[\tau + \delta]}{d\delta} \right\} \\
&= \lim_{\delta \rightarrow 0} \{-f(t, \tau + \delta) \cdot 1\} \\
&= -f(t, \tau) \quad (108)
\end{aligned}$$

and $\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\}$ is the expression for annualized instantaneous volatility, which I derive in section B.4 and denote $\omega(\tau)$, i.e.:

$$\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \tau + \delta) \right\} = \omega(\tau) \quad (109)$$

Substituting these results into equation 107 gives:

$$\lim_{\delta \rightarrow 0} d_1(t, \tau, \delta) = -\frac{f(t, \tau)}{\omega(\tau)} \quad (110)$$

and the final result for $\lim_{\delta \rightarrow 0} N[d_1(t, \tau, \delta)]$ is obtained substituting that result into equation 106, i.e.:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \Phi[d_1(t, \tau, \tau + \delta)] &= \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n \left[\frac{1}{\sqrt{2}} \cdot \left(-\frac{f(t, \tau)}{\omega(\tau)} \right) \right]^{2n+1}}{n! (2n+1)} \\
&= \frac{1}{2} + \frac{1}{2} \cdot \operatorname{erf} \left[-\frac{1}{\sqrt{2}} \frac{f(t, \tau)}{\omega(\tau)} \right] \\
&= \Phi \left[-\frac{f(t, \tau)}{\omega(\tau)} \right] \\
&= 1 - \Phi \left[\frac{f(t, \tau)}{\omega(\tau)} \right] \quad (111)
\end{aligned}$$

$$\mathbf{C.0.3} \quad \lim_{\delta \rightarrow 0} \left\{ \frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \frac{\mathbf{d}}{\mathbf{d}\delta} \Phi [d_1(t, \tau, \tau + \delta)] - \frac{\mathbf{d}}{\mathbf{d}\delta} \Phi [d_2(t, \tau, \tau + \delta)] \right\}$$

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

$$\begin{aligned} \frac{\mathbf{d}}{\mathbf{d}\delta} \Phi [d_1(t, \tau, \delta)] &= \frac{\mathbf{d}}{\mathbf{d}x} \Phi [x] \frac{\mathbf{d}}{\mathbf{d}\delta} d_1(t, \tau, \delta) \\ \langle \text{Chain rule} \rangle &: \left\langle x = d_1(t, \tau, \delta); \frac{\mathbf{d}}{\mathbf{d}x} \mathbf{N}[x] = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} x^2 \right] \right\rangle \\ &= \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(t, \tau, \delta)]^2 \right] \frac{\mathbf{d}}{\mathbf{d}\delta} d_1(t, \tau, \delta) \end{aligned} \quad (112)$$

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

$$\frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \frac{\mathbf{d}}{\mathbf{d}\delta} \Phi [d_1(t, \tau, \delta)] = \frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(\cdot)]^2 \right] \frac{\mathbf{d}}{\mathbf{d}\delta} d_1(\cdot) \quad (113)$$

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

$$\frac{\mathbf{d}}{\mathbf{d}\delta} \Phi [d_2(t, \tau, \delta)] = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_2(t, \tau, \delta)]^2 \right] \frac{\mathbf{d}}{\mathbf{d}\delta} d_2(t, \tau, \delta) \quad (114)$$

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

$$\begin{aligned} [d_2(t, \tau, \delta)]^2 &= [d_1(t, \tau, \delta) - \Sigma(\tau, \delta)]^2 \\ &= [d_1(\cdot)]^2 - 2d_1(\cdot) \Sigma(\tau, \tau + \delta) + [\Sigma(\tau, \delta)]^2 \\ &= [d_1(\cdot)]^2 - 2 \left(\frac{1}{\Sigma(\cdot)} \log \left[\frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \right] + \frac{1}{2} \Sigma(\cdot) \right) \Sigma(\cdot) - [\Sigma(\cdot)]^2 \\ &= [d_1(\cdot)]^2 - 2 \cdot \log \left[\frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \right] + [\Sigma(\cdot)]^2 - [\Sigma(\cdot)]^2 \\ &= [d_1(\cdot)]^2 - 2 \cdot \log \left[\frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \right] \end{aligned} \quad (115)$$

and therefore:

$$\begin{aligned} \exp \left[-\frac{1}{2} [d_2(t, \tau, \delta)]^2 \right] &= \exp \left[-\frac{1}{2} \left([d_1(\cdot)]^2 - 2 \cdot \log \left[\frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \right] \right) \right] \\ &= \exp \left[-\frac{1}{2} [d_1(\cdot)]^2 \right] \cdot \exp \left[\log \left(\frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \right) \right] \\ &= \frac{\mathbf{P}(t, \tau + \delta)}{\mathbf{P}(t, \tau)} \exp \left[-\frac{1}{2} [d_1(t, \tau, \delta)]^2 \right] \end{aligned} \quad (116)$$

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

$$\begin{aligned} \frac{\mathbf{d}}{\mathbf{d}\delta} d_2(t, \tau, \delta) &= \frac{\mathbf{d}}{\mathbf{d}\delta} [d_1(t, \tau, \delta) - \Sigma(\tau, \delta)] \\ &= \frac{\mathbf{d}}{\mathbf{d}\delta} d_1(t, \tau, \delta) - \frac{\mathbf{d}}{\mathbf{d}\delta} \Sigma(\tau, \delta) \end{aligned} \quad (117)$$

Substituting the results from equations 116 and 117 into equation 114 gives the following expression:

$$\begin{aligned}
\frac{d}{d\delta}\Phi [d_2(\cdot)] &= \frac{1}{\sqrt{2\pi}} \frac{P(t, \tau + \delta)}{P(t, \tau)} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \left[\frac{d}{d\delta}d_1(\cdot) - \frac{d}{d\delta}\Sigma(\tau, \delta)\right] \\
&= \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta}d_1(\cdot) \\
&\quad - \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta}\Sigma(\tau, \delta) \tag{118}
\end{aligned}$$

When calculating $U(t, \tau, \delta) = \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{d}{d\delta}\Phi [d_1(t, \tau, \delta)] - \frac{d}{d\delta}\Phi [d_2(t, \tau, \delta)]$, note that the expression $\frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta}d_1(\cdot)$ in equation 113 cancels identically with the first component from equation 118. Therefore:

$$U(t, \tau, \delta) = \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta}\Sigma(\cdot) \tag{119}$$

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

$$\begin{aligned}
\lim_{\delta \rightarrow 0} U(t, \tau, \delta) &= \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \frac{d}{d\delta}\Sigma(\tau, \delta) \right\} \\
\langle \text{Limit rules} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \right\} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \right\} \\
&\quad \times \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta}\Sigma(\tau, \delta) \right\} \tag{120}
\end{aligned}$$

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

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{P(t, \tau + \delta)}{P(t, \tau)} \right\} &= \frac{P(t, \tau)}{P(t, \tau)} \\
&= 1 \tag{121}
\end{aligned}$$

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

$$\exp[x] = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{122}$$

Substituting $x = -\frac{1}{2} [d_1(t, \tau, \delta)]^2$ into equation 122 and taking the limit gives:

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} [d_1(\cdot)]^2\right] \right\} &= \frac{1}{\sqrt{2\pi}} \lim_{\delta \rightarrow 0} \left\{ \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2} [d_1(\cdot)]^2\right)^n}{n!} \right\} \\
&= \frac{1}{\sqrt{2\pi}} \lim_{\delta \rightarrow 0} \left\{ \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n [d_1(\cdot)]^{2n}}{n!} \right\} \\
\langle \text{Limit rules} \rangle &= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n \lim_{\delta \rightarrow 0} \left\{ [d_1(\cdot)]^{2n} \right\}}{n!} \\
\langle \text{Limit rules} \rangle &= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n [\lim_{\delta \rightarrow 0} d_1(t, \tau, \delta)]^{2n}}{n!} \tag{123}
\end{aligned}$$

The result $\lim_{\delta \rightarrow 0} [d_1(t, \tau, \delta)] = -\frac{f(t, \tau)}{\omega(\tau)}$ is available from equation 110, and so

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} [d_1(t, \tau, \delta)]^2 \right] \right\} &= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n \left[-\frac{f(t, \tau)}{\omega(\tau)}\right]^{2n}}{n!} \\
&= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n \left(\left[-\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right)^n}{n!} \\
&= \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right)^n}{n!} \\
&= \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2 \right) \tag{124}
\end{aligned}$$

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

$$\begin{aligned}
\lim_{\delta \rightarrow 0} U(t, \tau, \delta) &= 1 \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2 \right) \cdot \omega(\tau) \\
&= \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2 \right) \tag{125}
\end{aligned}$$

C.1 Annualized instantaneous volatility

I define annualized instantaneous volatility $\omega(\tau)$ as the annualized limit of the option volatility expression from equation 92 and denote it $\omega(\tau)$, i.e.:

$$\begin{aligned}
\omega(\tau) &= \lim_{\delta \rightarrow 0} \left\{ \frac{1}{\delta} \Sigma(\tau, \delta) \right\} \\
\langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \delta) \right\}}{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \delta \right\}} \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \delta) \right\} \tag{126}
\end{aligned}$$

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

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

$$\begin{aligned}
\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \delta) \right\} &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \sqrt{[\Sigma(\tau, \delta)]^2} \right\} \\
\langle \text{Chain rule} \rangle &= \lim_{\delta \rightarrow 0} \left\{ \frac{d}{dx} \sqrt{x} \frac{d}{d\delta} [\Sigma(\tau, \delta)]^2 \right\} \\
&\left\langle x = [\Sigma(\tau, \tau + \delta)]^2; \frac{d}{dx} \sqrt{x} = \frac{1}{2} x^{-1/2} = \frac{1}{2\sqrt{x}} \right\rangle \\
&= \lim_{\delta \rightarrow 0} \left\{ \frac{\frac{d}{d\delta} [\Sigma(\tau, \delta)]^2}{2\Sigma(\tau, \delta)} \right\} \\
\langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \left[\frac{d}{d\delta} [\Sigma(\tau, \delta)]^2 \right] \right\}}{2 \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \delta) \right\}} \\
&= \frac{\lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \delta)]^2 \right\}}{2 \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \delta) \right\}} \tag{127}
\end{aligned}$$

Note that L'Hopital's rule has been used in the second-last step of equation 127 because in the middle line both $\frac{d}{d\delta} [\Sigma(\tau, \delta)]^2 = 2\Sigma(\tau, \delta) \frac{d}{d\delta} \Sigma(\tau, \delta)$ and $2\Sigma(\tau, \delta)$ would equal zero when evaluated at $\delta = 0$ (see equations 92 and 93), which would leave the limit undefined.

Re-arranging equation 127 gives:

$$\left[\lim_{\delta \rightarrow 0} \left\{ \frac{d}{d\delta} \Sigma(\tau, \delta) \right\} \right]^2 = \frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \delta)]^2 \right\} \tag{128}$$

and therefore:

$$\omega(\tau) = \sqrt{\frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \delta)]^2 \right\}} \tag{129}$$

For the calculation itself, $[\Sigma(\tau, \delta)]^2$ is obtained from equation 92 with the substitution of equation 93, i.e.:

$$\begin{aligned}
[\Sigma(\tau, \delta)]^2 &= \int_0^\tau \|\sigma(s, \delta)\|^2 ds \\
&= \int_0^\tau \left\| \int_s^{s+\delta} \sigma(u) du \right\|^2 ds \\
&= \int_0^\tau \left(\left[\int_s^{s+\delta} \sigma(u) du \right]' \left[\int_s^{s+\delta} \sigma(u) du \right] \right) ds \\
&= \int_0^\tau \left(\left[\int_0^\delta \sigma(s+v) dv \right]' \left[\int_0^\delta \sigma(s+v) dv \right] \right) ds \tag{130}
\end{aligned}$$

where I have re-expressed the square of the vector norm $\|a\|^2$ equivalently as the vector inner product $a'a$ in line 3, and have undertaken a change of variables $v = u - s$ in line 4. The first derivative may then be obtained directly using the chain rule and the

fundamental rule of calculus $\frac{d}{dx} \int_c^x y(t) dt = y(x)$, i.e.:

$$\begin{aligned}
\frac{d}{d\delta} [\Sigma(\tau, \delta)]^2 &= \frac{d}{d\delta} \int_0^\tau \left(\left[\int_0^\delta \sigma(s+v) dv \right]' \left[\int_0^\delta \sigma(s+v) dv \right] \right) ds \\
&= \int_0^\tau \frac{d}{d\delta} \left(\left[\int_0^\delta \sigma(s+v) dv \right]' \left[\int_0^\delta \sigma(s+v) dv \right] \right) ds \\
&= \int_0^\tau \left\{ \left[\frac{d}{d\delta} \int_0^\delta \sigma(s+v) dv \right]' \left[\int_0^\delta \sigma(s+v) dv \right] \right. \\
&\quad \left. + \left[\int_0^\delta \sigma(s+v) dv \right]' \left[\frac{d}{d\delta} \int_0^\delta \sigma(s+v) dv \right] \right\} ds \\
&= \int_0^\tau \left\{ [\sigma(s+\delta)]' \left[\int_0^\delta \sigma(s+v) dv \right] \right. \\
&\quad \left. + \left[\int_0^\delta \sigma(s+v) dv \right]' [\sigma(s+\delta)] \right\} ds \tag{131}
\end{aligned}$$

and the second derivative directly as:

$$\begin{aligned}
\frac{d^2}{d\delta^2} [\Sigma(\tau, \delta)]^2 &= \frac{d}{d\delta} \left[\frac{d}{d\delta} [\Sigma(\tau, \delta)]^2 \right] \\
&= \int_0^\tau \frac{d}{d\delta} \left\{ [\sigma(s+\delta)]' \left[\int_0^\delta \sigma(s+v) dv \right] \right. \\
&\quad \left. + \left[\int_0^\delta \sigma(s+v) dv \right]' [\sigma(s+\delta)] \right\} ds \\
&= \int_0^\tau \left\{ [\sigma(s+\delta)]' \left[\frac{d}{d\delta} \int_0^\delta \sigma(s+v) dv \right] \right. \\
&\quad + \left[\frac{d}{d\delta} \int_0^\delta \sigma(s+v) dv \right]' [\sigma(s+\delta)] \\
&\quad + \frac{d}{d\delta} [\sigma(s+\delta)]' \left[\int_0^\delta \sigma(s+v) dv \right] \\
&\quad \left. + \left[\int_0^\delta \sigma(s+v) dv \right]' \frac{d}{d\delta} [\sigma(s+\delta)] \right\} ds \\
&= \int_0^\tau \{ [\sigma(s+\delta)]' [\sigma(s+\delta)] \} ds + \int_0^\tau [\sigma(s+\delta)]' [\sigma(s+\delta)] ds \\
&\quad + \int_0^\tau U(s, \delta) ds \\
&= 2 \int_0^\tau [\sigma(s, \delta)]' [\sigma(s, \delta)] ds + \int_0^\tau U(s, \delta) ds \tag{132}
\end{aligned}$$

where I have used the notation $\int_0^\tau U(s, \delta) ds$ to avoid the unnecessary repetition of the terms in lines 6 and 7, given they become zero when $\lim_{\delta \rightarrow 0} \{ \cdot \}$ is applied (i.e. the

upper and lower limits of the integrals $\int_0^\delta \sigma(s+v)dv$ in $U(s, \delta)$ become equal, while $\frac{d}{d\delta} [\sigma(s+\delta)]$ remains well-defined).

Therefore the final expression of $\omega(\tau)$ is

$$\begin{aligned}\omega(\tau) &= \sqrt{\frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ \frac{d^2}{d\delta^2} [\Sigma(\tau, \delta)]^2 \right\}} \\ &= \sqrt{\frac{1}{2} \cdot \lim_{\delta \rightarrow 0} \left\{ 2 \int_0^\tau [\sigma(s+\delta)]' [\sigma(s+\delta)] ds \right\}} \\ &= \sqrt{\int_0^\tau [\sigma(s)]' [\sigma(s)] ds}\end{aligned}\tag{133}$$

Finally, with reference to equation 83 in section B.1, the volatility function $[\sigma(s)]$ may be expressed as $[\sigma(s)] = \sigma[\mathbb{1}(b_0, \tilde{\kappa}, s)]$ for the shadow-GATSM. That expression therefore produces the final result as given in section 4.3, i.e.:

$$\omega(\tau) = \sqrt{\int_0^\tau [\mathbb{1}(b_0, \tilde{\kappa}, s)]' \sigma' \sigma [\mathbb{1}(b_0, \tilde{\kappa}, s)] ds}\tag{134}$$

D CAB-GATSM option effect: forward measure derivation

This appendix provide the details for the option effect derivation summarized in section 4.3.

First, it is a standard result for any term structure model that the expected value of the short rate under the risk-adjusted $t+\tau$ forward measure equals the forward rate.⁴⁰ Hence, for any shadow term structure model:

$$\tilde{\mathbb{E}}_{t+\tau} \{r(t+\tau) | \mathcal{F}_t\} = f(t, \tau)\tag{135}$$

where \mathcal{F}_t represents the information set available at time t , and expectations are formed conditional on \mathcal{F}_t .

The distribution of $\tilde{\mathbb{E}}_{t+\tau} \{r(t+\tau) | \mathcal{F}_t\}$ requires the distribution of $x(t+\tau)$ under the risk-adjusted measure. That result can be established more directly from the solution of the stochastic differential equation for $x(t)$, i.e. $dx(t) = \tilde{\kappa} [\tilde{\theta} - x(t)] dt + \sigma dW(t)$ as given in equation 35. Meucci (2010) p. 3, for example, gives the following solution for $x(t+\tau)$ conditional on $x(t)$:⁴¹

$$x(t+\tau) = [I - \exp(-\tilde{\kappa}\tau)] \tilde{\theta} + \exp(-\tilde{\kappa}\tau) x(t) + \int_0^\tau \sigma \exp(-\tilde{\kappa}s) dW(s)\tag{136}$$

⁴⁰See, for example, Filipović (2009) p. 107, James and Webber (2000) p. 99-100, or Klebaner (2005) p. 337 which discusses the Heath, Jarrow, and Morton (1992) framework under the forward measure. These references all use time and time of maturity notation $\tilde{\mathbb{E}}_T \{r(T) | \mathcal{F}_t\} = f(t, T)$, which I express equivalently using my time and time to maturity notation.

⁴¹I have made the substitution $s = v - t$ to convert the integral $\int_t^{t+\tau} \sigma \exp(-\tilde{\kappa}[v-t]) dW(v)$ from Meucci (2010) to my integral expression in equation 136.

Therefore, the covariance matrix for $x(t + \tau)$ may be expressed as:

$$\text{cov}[x(t + \tau)] = \int_0^\tau [\exp(-\tilde{\kappa}s)]' \sigma' \sigma [\exp(-\tilde{\kappa}s)] ds \quad (137)$$

and the variance of $r(t + \tau)$ is:

$$\begin{aligned} [\omega(\tau)]^2 &= \text{var}[r(t + \tau)] \\ &= b_0' \text{cov}[x(t + \tau)] b_0 \\ &= b_0' \left[\int_0^\tau [\exp(-\tilde{\kappa}s)]' \sigma' \sigma [\exp(-\tilde{\kappa}s)] ds \right] b_0 \\ &= \int_0^\tau [\mathbb{1}(b_0, \tilde{\kappa}, s)]' \sigma' \sigma [\mathbb{1}(b_0, \tilde{\kappa}, s)] ds \end{aligned} \quad (138)$$

Combining the mean and variance results gives the distribution of risk-adjusted short rates under the $t + \tau$ forward measure for the generic shadow-GATSM as the following normal distribution:

$$r(t + \tau) \sim N \{f(t, \tau), [\omega(\tau)]^2\} \quad (139)$$

Using that distribution, the option effect for the CAB-GATSM may then be obtained as the expected value of $\max\{-r(t + \tau), 0\}$ under the risk-adjusted $t + \tau$ forward measure. Specifically, repeating the following result from the main text:

$$\begin{aligned} z(t, \tau) &= \tilde{\mathbb{E}}_{t+\tau} [\max\{-r(t + \tau), 0\}] \\ &= - \int_{-\infty}^0 r(t + \tau) \cdot \text{Prob}[r(t + \tau)] dr(t + \tau) \end{aligned} \quad (140)$$

and substituting $y \equiv r(t + \tau)$ for notational convenience results in the following expression to solve:

$$z(t, \tau) = \int_{-\infty}^0 -y \cdot \frac{1}{\omega(\tau)} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{y - f(t, \tau)}{\omega(\tau)}\right]^2\right) dy \quad (141)$$

I believe the indefinite integral may be solved via substitution and integration by parts, but I have simply used Mathematica to obtain that result. The expression for the definite integral is then:

$$\begin{aligned} z(t, \tau) &= \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{y - f(t, \tau)}{\omega(\tau)}\right]^2\right) - f(t, \tau) \cdot \frac{1}{2} \text{erf}\left[\frac{1}{\sqrt{2}} \frac{y - f(t, \tau)}{\omega(\tau)}\right] \Bigg|_{-\infty}^0 \\ &= \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{f(t, \tau)}{\omega(\tau)}\right]^2\right) - 0 \\ &\quad - f(t, \tau) \cdot \left(\frac{1}{2} \text{erf}\left[\frac{1}{\sqrt{2}} \frac{-f(t, \tau)}{\omega(\tau)}\right] - \frac{1}{2}\right) \\ &= \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{y - f(t, \tau)}{\omega(\tau)}\right]^2\right) - f(t, \tau) \cdot \left(\frac{1}{2} - \frac{1}{2} \text{erf}\left[\frac{1}{\sqrt{2}} \frac{f(t, \tau)}{\omega(\tau)}\right]\right) \\ &= -f(t, \tau) \cdot \left(1 - \Phi\left[\frac{f(t, \tau)}{\omega(\tau)}\right]\right) + \omega(\tau) \cdot \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{f(t, \tau)}{\omega(\tau)}\right)^2\right] \end{aligned} \quad (142)$$

where $\text{erf}[\cdot]$ is the error function, and the final substitution follows from the definition of the unit cumulative normal distribution in terms of the error function, i.e.:

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

so therefore:

$$\begin{aligned} 1 - \Phi[x] &= 1 - \left[\frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{x}{\sqrt{2}}\right) \right] \\ &= \frac{1}{2} - \frac{1}{2} \text{erf}\left(\frac{x}{\sqrt{2}}\right) \end{aligned} \tag{143}$$

E Applying the iterated extended Kalman filter

This appendix contains the details underlying the application of the IEKF to CAB-GATSM specification in general. Appendix E.1 provides an overview of the IEKF algorithm and estimation via maximum likelihood, appendix E.2 presents further details underlying each of the individuals steps of the algorithm, and section E.3 briefly describes the parameter restrictions I have used for the estimations in sections 5.2 to 5.4.

E.1 IEKF and CAB-GATSM estimation overview

I adopt the IEKF algorithm and notation from Simon (2006). Hence, in what follows, x_t^- and P_t^- are the prior estimates of the state vector and its covariance at the integer time step t , $x_{t,i}^+$ and $P_{t,i}^+$ are the posterior estimates at iteration i and t , and x_t^+ and P_t^+ are the final posterior estimates at t .

The four steps I summarize below follow the IEKF summary from Simon (2006) p. 412, and they are analogous to the standard (linear) Kalman filter with the following exception: the constant $K \times N$ matrix H of coefficients that relates observations to the state in the standard Kalman filter is replaced with $H(x_{t,i}^+, \mathbb{A})$. The latter is obtained as the partial derivative of $\mathbb{R}[x(t), \mathbb{A}]$ with respect to $x(t)$ evaluated at $x_{t,i}^+$, and a new evaluation of $H(x_{t,i}^+, \mathbb{A})$ is required at each iteration i and time step t . However, it conveniently turns out (from the perspective of empirical implementation), as I detail in appendix E.2.3, that evaluating $H(x_{t,i}^+, \mathbb{A})$ requires no more numerical evaluations than are already required to obtain $\mathbb{R}(x_{t,i}^+, \mathbb{A})$.

E.1.1 IEKF steps

1. Initialization:

$$x_0^+ = \theta \tag{144a}$$

$$P_0^+ = V\Theta(\infty)V^{-1} \tag{144b}$$

where $\kappa = V\kappa_D V^{-1}$ with $\kappa_D = \text{diag}[\kappa_1, \dots, \kappa_i, \dots, \kappa_N]$ and:

$$[\Theta(\infty)]_{ij} = \frac{1}{\kappa_i + \kappa_j} \tag{145}$$

2. Prior estimates:

$$x_{t+1}^- = [I - F] \theta + F x_t^+ \quad (146a)$$

$$P_{t+1}^- = F P_t^+ F' + \Theta(\tau) \quad (146b)$$

where:

$$[\Theta(\Delta t)]_{ij} = \frac{1}{\kappa_i + \kappa_j} [1 - \exp(-\{\kappa_i + \kappa_j\} \Delta t)] \quad (147)$$

3. Measurement updates:

$$\text{Set} : x_{t,0}^+ = x_t^- \quad (148a)$$

$$\text{ITERATE} : \text{on } i \quad (148b)$$

$$H_{t,i} = \left. \frac{\partial}{\partial x(t)} \mathbb{R}[x(t), \mathbb{A}, \tau_k] \right|_{x(t)=x_{t,i}^+} \quad (148c)$$

$$M_{t,i} = H_{t,i} P_t^- H_{t,i}' + \Omega_\eta \quad (148d)$$

$$K_{t,i} = P_t^- H_{t,i}' M_{t,i}^{-1} \quad (148e)$$

$$\eta_{t,i} = \mathbb{R}_t - \mathbb{R}(x_{t,i}^+, \mathbb{A}) - H_{t,i} (x_t^- - x_{t,i}^+) \quad (148f)$$

$$x_{t,i+1}^+ = x_{t,i}^- + K_{t,i} \eta_{t,i} \quad (148g)$$

$$\text{EXIT} : \text{at } \max(i) \text{ or a tolerance for } [x_{t,i+1}^+ - x_{t,i}^+] \quad (148h)$$

4. Posterior estimates:

$$x_t^+ = x_{t,i+1}^+ \quad (149a)$$

$$P_t^+ = (I - K_{t,i} H_{t,i}) P_{t,i}^+ \quad (149b)$$

Note that the EKF is nested within the IEKF procedure outlined in Simon (2006), because the zeroth iteration is set to the posterior estimate. Therefore, EKF results may be obtained by choosing not to iterate after the first update step; i.e. with $x_{t,0}^+ = x_t^-$, $x_t^- - x_{t,0}^+ = 0$, and so $x_t^+ = x_t^- + K_t [\mathbb{R}_t - \mathbb{R}(x_t^-, \mathbb{A})]$. However, some initial results using the EKF in related work proved suspiciously variable (see footnote , which ultimately led me to iterate the IEKF until all elements of $|x_{t,i+1}^+ - x_{t,i+1}^-|$ were each less than 1e-8. I illustrate and discuss this issue in section 5.4.

E.1.2 Parameter estimation

Each update of the IEKF for a given set of parameters \mathbb{A} provides a log-likelihood value as follows:

$$\log L(\{\mathbb{R}_1, \dots, \mathbb{R}_T\}, \mathbb{A}) = -\frac{1}{2} \sum_{t=1}^T K \log(2\pi) + \log M_{t,i} + \eta_{t,i}' M_{t,i}^{-1} \eta_{t,i} \quad (150)$$

and I use the Nelder-Mead algorithm (from the “fminsearch” function in the MatLab Optimization Toolbox) to find the parameter set \mathbb{A} for the given model specifications (in sections 5.3 to 5.5) that maximize the log-likelihood value. The tolerance I use to indicate convergence to the maximum value is 0.01 log-likelihood points.

E.2 Iterated extended Kalman filter steps

E.2.1 Initialization

The state variable vector and its covariance matrix should be initialized at their unconditional expected values, which for the CAB-GATSM are respectively $\mathbb{E}_0[x(t)]$ and $\mathbb{E}_0(\{x(t) - \mathbb{E}_0[x(t)]\}\{x(t) - \mathbb{E}_0[x(t)]\}')$. Calculating those quantities requires the solution for the dynamic process of the state variable vector $x(t)$ under the physical measure. That solution may be obtained by directly solving the stochastic differential equation, which is analogous to the solution already presented in appendix B.5 (or see, for example, Klebaner (2005) p. 127-28). Hence:

$$x(t) = [I - \exp(-\kappa t)]\theta + \exp(-\kappa t)x(0) + \int_0^t \sigma \exp(-\kappa v) dW(v) \quad (151)$$

where $x(0)$ is the initial value of $x(t)$, i.e. at time $t = 0$. The expected value of $x(t)$ as at time $t = 0$ is therefore:

$$\mathbb{E}_0[x(t)] = [I - \exp(-\kappa t)]\theta + \exp(-\kappa t)x(0) \quad (152)$$

and so the unconditional expectation of $x(t)$, i.e. with $t = 0$ in the infinite past relative to the present time t , is:

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E}_0[x(t)] &= [I - 0]\theta + 0x(0) \\ &= \theta \end{aligned} \quad (153)$$

Hence, $x_0^+ = \theta$ is the appropriate initial value for the IEKF.

The covariance matrix for $x(t)$ at time t is:

$$\begin{aligned} \Omega_{x,t}[x(t)] &= \int_0^t [\exp(-\kappa v)]' \sigma' \sigma \exp(-\kappa v) ds \\ &= \int_0^t [\exp(-V\kappa_D V^{-1}v)]' \Omega [\exp(-V\kappa_D V^{-1}v)] dv \\ &= \int_0^t [V \exp(-\kappa_D v) V^{-1}]' \Omega [V \exp(-\kappa_D v) V^{-1}] dv \\ &= V \left\{ \int_0^t [\exp(-\kappa_D v)] V'^{-1} \Omega V [\exp(-\kappa_D v)] dv \right\} V^{-1} \\ &= V\Theta(t)V^{-1} \end{aligned} \quad (154)$$

where $\kappa = V\kappa_D V^{-1}$ is the eigensystem decomposition of the $N \times N$ mean-reversion matrix κ ,⁴² and Θ is notation for the $N \times N$ matrix function resulting from the definite integral. Making the substitution $U = V'^{-1}\Omega V$ for notational convenience, the

⁴²Note that an allowance for repeated eigenvalues is not required because the only constraint I impose on κ is that the real part of the eigenvalues be non-negative; see appendix E.3. Conversely, the three-factor model in section 5.4 constrains the risk-adjusted mean-reversion matrix $\tilde{\kappa}$ to have one eigenvalue equal to 0 and two repeated eigenvalues equal to ϕ . However, repeated eigenvalues in κ could readily be allowed for using the Jordan decomposition for κ rather the eigensystem decomposition, and that may be appropriate for numerical stability in some estimations.

elements for Θ may be calculated as follows:

$$\begin{aligned}
[\Theta(t)]_{ij} &= \int_0^t [\exp(-\kappa_i v)] U_{ij} [\exp(-\kappa_j v)] dv \\
&= \int_0^t U_{ij} \exp(-[\kappa_i + \kappa_j] v) dv \\
&= -\frac{1}{\kappa_i + \kappa_j} \exp(-[\kappa_i + \kappa_j] s) \Big|_0^t \\
&= \frac{1}{\kappa_i + \kappa_j} [1 - \exp(-[\kappa_i + \kappa_j] t)]
\end{aligned} \tag{155}$$

Hence, the unconditional expectation of $x(t)$ is:

$$\lim_{t \rightarrow \infty} \Omega_{x,0} [x(t)] = V \Theta(\infty) V^{-1} \tag{156}$$

where:

$$\begin{aligned}
[\Theta(\infty)]_{ij} &= \lim_{t \rightarrow \infty} [\Theta(t)]_{ij} \\
&= \frac{1}{\kappa_i + \kappa_j}
\end{aligned} \tag{157}$$

E.2.2 Prior estimates

The prior estimate is the expected value of the state variable vector for the next time step, $x(t + \Delta t)$ conditional on its current value $x(t)$. The required expression may be obtained via the conditional expectation result under the physical measure, which is analogous to the solution already presented in appendix B.5. Specifically:

$$x(t + \tau) = [I - \exp(-\kappa\tau)] \theta + \exp(-\kappa\tau) x(t) + \int_0^\tau \sigma \exp(-\kappa s) dW(s) \tag{158}$$

and then substituting the time increment Δt for τ and taking expectations under the risk-adjusted measures as at time t gives:

$$\tilde{\mathbb{E}}_t [x(t + \Delta t)] = [I - \exp(-\tilde{\kappa}\Delta t)] \tilde{\theta} + \exp(-\tilde{\kappa}\Delta t) x(t) \tag{159}$$

Allowing for innovations to the state equation, that result may be re-expressed as:

$$x(t + \Delta t) = [I - F] \tilde{\theta} + Fx(t) + \varepsilon(t + \Delta t) \tag{160}$$

where:

$$F = \exp(-\kappa\Delta t) \tag{161}$$

and $\varepsilon(t + \Delta t)$ is an $N \times 1$ vector of innovations with the $N \times N$ covariance matrix:

$$\begin{aligned}
\Omega_\varepsilon &= \int_0^\tau \exp(-\kappa s) \sigma' \sigma \exp(-\kappa s) ds \\
&= V \Theta(\tau) V^{-1}
\end{aligned} \tag{162}$$

and the expression $V\Theta(\tau)V^{-1}$ results from an eigensystem decomposition analogous to that in appendix C.1.1. The elements of $\Theta(\tau)$ are:

$$[\Theta(\tau)]_{ij} = \frac{1}{\kappa_i + \kappa_j} [1 - \exp(-[\kappa_i + \kappa_j]\tau)] \quad (163)$$

Finally, as in the main text, I further re-express equation 160 using the subscript t to represent an integer index of time increments Δt over the sample period from time t to T , i.e.:

$$x_{t+1} = [I - F]\tilde{\theta} + Fx_t + \varepsilon_{t+1} \quad (164)$$

E.2.3 Measurement updates

CAB-GATSM interest rates There is no closed-form analytic solution available for CAB-GATSM interest rates $\mathbb{R}(x_t, \mathbb{A}, \tau_k)$ required in the measurement equation, because the solution involves the integral of the standard cumulative normal distribution. However, the integral can be approximated to arbitrary precision using numerical methods. A particularly convenient form is a rectangular rule with constant time-to-maturity increments $\Delta\tau$ and end-increment function values, i.e.:

$$\mathbb{R}(x_t, \mathbb{A}, \tau_k) \simeq \frac{1}{\tau_k} \left(\sum_{i=1}^{I_k} \underline{f}(x_t, \mathbb{A}, i\Delta\tau) \Delta\tau \right) \quad (165)$$

where $I_k = \tau_k/\Delta\tau$, and $\underline{f}(x_t, \mathbb{A}, \Delta\tau), \underline{f}(x_t, \mathbb{A}, 2\Delta\tau), \dots, \underline{f}(x_t, \mathbb{A}, I_k\Delta\tau)$ is a sequence of forward rates $\underline{f}(x_t, \mathbb{A}, u)$ for times to maturity $u = i\Delta\tau$ and:

$$\underline{f}(x_t, \mathbb{A}, u) = f(x_t, \mathbb{A}, u) \cdot \Phi \left[\frac{f(x_t, \mathbb{A}, u)}{\omega(\mathbb{A}, u)} \right] + \omega(\mathbb{A}, u) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(x_t, \mathbb{A}, u)}{\omega(\mathbb{A}, u)} \right]^2 \right) \quad (166)$$

Equation 165 turns out to be an arithmetic mean, i.e.:

$$\begin{aligned} \mathbb{R}(x_t, \mathbb{A}, \tau_k) &\simeq \frac{1}{\tau_k} \left(\sum_{i=1}^{I_k} \underline{f}(x_t, \mathbb{A}, i\Delta\tau) \Delta\tau \right) \\ &= \frac{1}{\tau_k} \Delta\tau \sum_{i=1}^{I_k} \underline{f}(x_t, \mathbb{A}, i\Delta\tau) \\ &= \frac{1}{\tau_k} \frac{\tau_k}{I_k} \sum_{i=1}^{I_k} \underline{f}(x_t, \mathbb{A}, i\Delta\tau) \\ &= \frac{1}{I_k} \sum_{i=0}^{I_k} \underline{f}(x_t, \mathbb{A}, i\Delta\tau) \\ &= \text{mean} \{ \underline{f}(t, \mathbb{A}, \Delta\tau), \underline{f}(t, \mathbb{A}, 2\Delta\tau), \dots, \underline{f}(t, \mathbb{A}, I_k\Delta\tau) \} \end{aligned} \quad (167)$$

Note that, while equations 167 and 166 provide convenient and transparent notation, the values of $\mathbb{R}(x_t, \mathbb{A}, \tau_k)$ should not be calculated independently for each time to maturity as written. Doing so would result in a very inefficient replication of numerical evaluations for $f(x_t, \mathbb{A}, u)$, $\omega(\mathbb{A}, u)$, and associated calculations for values of u

that are common to all maturities. Rather, I first calculate a single series of shadow-GATSM forward rates and annualized option volatilities out to the longest time to maturity τ_K ; i.e. $f(x_t, \mathbb{A}, \Delta\tau), f(x_t, \mathbb{A}, 2\Delta\tau), \dots, f(x_t, \mathbb{A}, \tau_K)$ and $\omega(\mathbb{A}, \Delta\tau), \omega(\mathbb{A}, 2\Delta\tau), \dots, \omega(\mathbb{A}, \tau_K)$. I use those results to obtain the associated CAB-GATSM forward rates $\underline{f}(x_t, \mathbb{A}, \Delta\tau), \underline{f}(x_t, \mathbb{A}, 2\Delta\tau), \dots, \underline{f}(x_t, \mathbb{A}, \tau_K)$ for each time to maturity, and then obtain all values of $\underline{R}(x_t, \mathbb{A}, \tau_k)$ at once by calculating the respective means of that series up to the elements $I_1 = \tau_1/\Delta\tau, I_2 = \tau_2/\Delta\tau, \dots, I_K = \tau_K/\Delta\tau$.

CAB-GATSM Jacobian At any given iteration i , the Jacobian $H_{t,i}$ could be evaluated numerically at the cost of numerically evaluating the combinations of the elements $\underline{R}[x_i(t) + \Delta x, \mathbb{A}, \tau_k]$ and $\underline{R}[x_j(t) + \Delta x, \mathbb{A}, \tau_k]$ in addition to the central value $\underline{R}[x(t), \mathbb{A}, \tau_k]$. However, it turns out that the closed-form analytic expression for forward rates in the CAB-GATSM framework can again be exploited, which leads to an expression for the Jacobian that requires no numerical evaluations in addition to those already undertaken to obtain $\underline{R}[x(t), \mathbb{A}, \tau_k]$. Specifically, I begin with the definition of the element of the Jacobian $H_{t,i}$ corresponding to the interest rate $\underline{R}[x(t), \mathbb{A}, \tau_k]$, and then re-express the latter in terms of forward rates, i.e.:

$$\begin{aligned} H_{t,i} &= \left. \frac{\partial}{\partial x(t)} \underline{R}[x(t), \mathbb{A}, \tau_k] \right|_{x(t)=x_{t,i}^+} \\ &= \left. \frac{\partial}{\partial x(t)} \left(\frac{1}{\tau_k} \int_0^{\tau_k} \underline{f}[x(t), \mathbb{A}, u] \, du \right) \right|_{x(t)=x_{t,i}^+} \\ &= \left. \frac{1}{\tau_k} \int_0^{\tau_k} \frac{\partial}{\partial x(t)} \underline{f}[x(t), \mathbb{A}, u] \, du \right|_{x(t)=x_{t,i}^+} \end{aligned} \quad (168)$$

Omitting the dependence on the parameter set \mathbb{A} for notational convenience, the CAB-GATSM forward rate expression $\underline{f}[x(t), u]$ is as follows:

$$\underline{f}[x(t), u] = f(t, \tau) \cdot \Phi \left[\frac{f(t, u)}{\omega(u)} \right] + \omega(u) \cdot \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, u)}{\omega(u)} \right]^2 \right) \quad (169)$$

and the partial differential of $\underline{f}[x(t), u]$ with respect to $x(t)$ is:

$$\begin{aligned} \frac{\partial}{\partial x(t)} \underline{f}[x(t), u] &= \frac{\partial}{\partial x_n(t)} f[x(t), u] \cdot \Phi \left[\frac{f[x(t), u]}{\omega(u)} \right] \\ &\quad + f[x(t), u] \cdot \frac{\partial}{\partial x(t)} \Phi \left[\frac{f[x(t), u]}{\omega(u)} \right] \\ &\quad + \omega(u) \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{\partial}{\partial x(t)} \exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \end{aligned} \quad (170)$$

I evaluate the result for each line in turn. Hence, the partial differential of $f[x(t), u]$ with respect to $x(t)$ is:

$$\begin{aligned} \frac{\partial}{\partial x(t)} f[x(t), u] &= \frac{\partial}{\partial x(t)} \{ \text{VE}(x(t), u) + [\mathfrak{f}(b_0, \tilde{\kappa}, u)]' x(t) \} \\ &= [\mathfrak{f}(b_0, \tilde{\kappa}, u)]' \end{aligned} \quad (171)$$

the partial differential of $\Phi \left[\frac{f[x(t), u]}{\omega(u)} \right]$ with respect to $x(t)$ is:

$$\begin{aligned}
\frac{\partial}{\partial x(t)} \Phi [\cdot] &= \frac{\partial}{\partial y} \Phi [y] \frac{\partial}{\partial x(t)} \left[\frac{f[x(t), u]}{\omega(u)} \right] \\
\langle \text{Chain rule} \rangle &: \left\langle y = \frac{f[x(t), u]}{\omega(u)}; \frac{\partial}{\partial y} \Phi [y] = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} y^2 \right) \right\rangle \\
&= \frac{1}{\sqrt{2\pi}} \cdot \exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \cdot \frac{\partial}{\partial x(t)} f[x(t), u] \cdot \frac{1}{\omega(u)} \\
&= \frac{1}{\sqrt{2\pi}} \cdot \exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \cdot \frac{1}{\omega(u)} \cdot [\mathfrak{I}(b_0, \tilde{\kappa}, u)]' \quad (172)
\end{aligned}$$

and the partial differential of $\exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right)$ with respect to $x(t)$ is:

$$\begin{aligned}
\frac{\partial}{\partial x(t)} \exp(\cdot) &= \frac{\partial}{\partial x(t)} \exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \\
\langle \text{Chain rule} \rangle &= \frac{\partial}{\partial y} \exp(y) \frac{\partial}{\partial x_n(t)} \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \\
&\left\langle y = -\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2; \frac{\partial}{\partial y} \exp(y) = \exp(y) \right\rangle \\
\langle \text{Chain rule} \rangle &= \exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \frac{\partial}{\partial z} \left(-\frac{1}{2} z^2 \right) \frac{\partial}{\partial x_n(t)} \left[\frac{f[x(t), u]}{\omega(u)} \right] \\
&\left\langle z = \frac{f[x(t), u]}{\omega(u)}; \frac{\partial}{\partial y} \left(-\frac{1}{2} z^2 \right) = -z \right\rangle \\
&= -\exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \frac{f[x(t), u]}{\omega(u)} \cdot \frac{1}{\omega(u)} \cdot \frac{\partial}{\partial x(t)} f[x(t), u] \\
&= -\exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \cdot \frac{f[x(t), u]}{[\omega(u)]^2} \cdot [\mathfrak{I}(b_0, \tilde{\kappa}, u)]' \quad (173)
\end{aligned}$$

Therefore, the final expression for the partial differential of $\mathfrak{f}[x(t), u]$ with respect to $x_n(t)$ is:

$$\begin{aligned}
\frac{\partial}{\partial x(t)} \mathfrak{f}(t, u) &= [\mathfrak{I}(b_0, \tilde{\kappa}, u)]' \cdot \Phi \left[\frac{f[x(t), u]}{\omega(u)} \right] \\
&+ f[x(t), u] \frac{1}{\sqrt{2\pi}} \cdot \exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \cdot \frac{1}{\omega(u)} \cdot [\mathfrak{I}(b_0, \tilde{\kappa}, u)]' \\
&- \omega(u) \cdot \exp \left(-\frac{1}{2} \left[\frac{f[x(t), u]}{\omega(u)} \right]^2 \right) \cdot \frac{f[x(t), u]}{[\omega(u)]^2} \cdot [\mathfrak{I}(b_0, \tilde{\kappa}, u)]' \\
&= [\mathfrak{I}(b_0, \tilde{\kappa}, u)]' \cdot \Phi \left[\frac{f[x(t), u]}{\omega(u)} \right] \quad (174)
\end{aligned}$$

Substituting the result for $\frac{\partial}{\partial x(t)}\underline{f}(t, u)$ into equation 168 then gives:

$$\begin{aligned} H_{t,i} &= \frac{1}{\tau_k} \int_0^{\tau_k} [\mathfrak{h}(b_0, \tilde{\kappa}, u)]' \cdot \Phi \left[\frac{\underline{f}(x(t), u)}{\omega(u)} \right] du \Big|_{x(t)=x_{t,i}^+} \\ &= \frac{1}{\tau_k} \int_0^{\tau_k} [\mathfrak{h}(b_0, \tilde{\kappa}, u)]' \cdot \Phi \left[\frac{\underline{f}(x_{t,i}^+, u)}{\omega(u)} \right] du \end{aligned} \quad (175)$$

Re-inserting the dependence on the parameter set \mathbb{A} , it is clear that the numerical evaluations of $\Phi \left[\frac{\underline{f}(x_{t,i}^+, \mathbb{A}, u)}{\omega(u)} \right]$ are already available from the evaluation of CAB-GATSM forward rates required to evaluate interest rates $\mathbb{R}[x(t), \mathbb{A}, \tau_k]$.

E.2.4 Posterior estimates

For the state variable vector, the final posterior estimates x_t^+ are set to the most recent iteration of the IEKF, i.e. $x_{t,i+1}^+$. The associated final posterior estimate for the covariance matrix P_t^+ is calculated using the prevailing values of $K_{t,i}$ and $H_{t,i}$, which therefore ensures that x_t^+ and P_t^+ are based on values from the same iteration.

E.3 Parameter constraints

For the physical mean-reversion matrix κ , I impose the constraints that each eigenvalue of κ always has a positive real part. If a realization during estimation is negative, I restrict the real part to 1e-6 while retaining the original eigenvectors. I also ensure that each standard deviation parameter is positive by using the standard mechanism $\sigma_{ij} = \sqrt{\sigma_{ij}^{*2}} > 0$, where σ_{ij}^* is the actual parameter estimated, and that each innovation correlation parameter ρ_{ij} is maintained between -1 and 1 using the standard mechanism:⁴³

$$-1 < \rho_{ij} = \frac{\rho_{ij}^*}{1 + |\rho_{ij}^*|} < 1 \quad (176)$$

where ρ_{ij}^* is the actual parameter estimated. Standard errors for the final parameter estimates (with σ_{ij}^* and ρ_{ij}^* converted back to σ_{ij} and ρ_{ij}) are obtained using a numerical Hessian evaluated around the final parameter estimates.

F CAB-GATSM(2) specification

With $b_0 = [1, 1]'$ and $\tilde{\kappa} = \text{diag}[\tilde{\kappa}_1, \tilde{\kappa}_2]$, the shadow-GATSM(2) forward rate factor loading vector function $[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]$ becomes:

$$\begin{aligned} [\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]' &= b_0' \exp(-\tilde{\kappa}\tau) \\ &= [1, 1] \exp \left(- \begin{bmatrix} \tilde{\kappa}_1 & 0 \\ 0 & \tilde{\kappa}_2 \end{bmatrix} \tau \right) \\ &= [1, 1] \begin{bmatrix} \exp(-\tilde{\kappa}_1\tau) & 0 \\ 0 & \exp(-\tilde{\kappa}_2\tau) \end{bmatrix} \\ &= [\exp(-\tilde{\kappa}_1\tau), \exp(-\tilde{\kappa}_2\tau)] \end{aligned} \quad (177)$$

⁴³See, for example, 146-47 for discussion on these mechanisms.

The expected path of the shadow short rate is therefore:

$$\begin{aligned}
\tilde{\mathbb{E}}_t [r(t + \tau)] &= a_0 + b'_0 \exp(-\tilde{\kappa}\tau) x(t) \\
&= a_0 + [\exp(-\tilde{\kappa}_1\tau), \exp(-\tilde{\kappa}_2\tau)] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \\
&= x_1(t) \cdot \exp(-\tilde{\kappa}_1\tau) + x_2(t) \cdot \exp(-\tilde{\kappa}_2\tau)
\end{aligned} \tag{178}$$

With the volatility matrix:

$$\sigma' = \begin{bmatrix} \sigma_1 & 0 \\ \rho_{12}\sigma_2 & \sigma_2\sqrt{1-\rho_{12}^2} \end{bmatrix} \tag{179}$$

the forward rate volatility loadings are defined as $\sigma(\tau) = \sigma[\mathbb{1}(b_0, \tilde{\kappa}, \tau)]$ and so the volatility effect is:

$$\begin{aligned}
\text{VE}(\tau) &= \int_0^\tau [\mathbb{1}(b_0, \tilde{\kappa}, \tau - s)]' \sigma' \left[\int_s^\tau \sigma[\mathbb{1}(b_0, \tilde{\kappa}, s - u)] du \right] ds \\
&= \int_0^\tau \left\{ [\exp(-\tilde{\kappa}_1\{\tau - s\}), \exp(-\tilde{\kappa}_2\{\tau - s\})] \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right. \\
&\quad \left. \times \left[\int_s^\tau \begin{bmatrix} \exp(-\tilde{\kappa}_1\{u - s\}) \\ \exp(-\tilde{\kappa}_2\{u - s\}) \end{bmatrix} du \right] ds \right\} \\
&= \sigma_1^2 \cdot \text{L/L} + \sigma_2^2 \cdot \text{S/S} + \rho_{12}\sigma_1\sigma_2 \cdot \text{L/S} + \rho_{12}\sigma_1\sigma_2 \cdot \text{S/L} \\
&= \sigma_1^2 \cdot \text{L/L} + \sigma_2^2 \cdot \text{S/S} + \rho_{12}\sigma_1\sigma_2 \cdot (\text{L/S} + \text{S/L})
\end{aligned} \tag{180}$$

where:

$$\text{L/L} = \int_0^\tau \exp(-\tilde{\kappa}_1[\tau - s]) \left(\int_s^\tau \exp(-\tilde{\kappa}_1[u - s]) du \right) ds \tag{181a}$$

$$\text{S/S} = \int_0^\tau \exp(-\tilde{\kappa}_2[\tau - s]) \left(\int_s^\tau \exp(-\tilde{\kappa}_2[u - s]) du \right) ds \tag{181b}$$

$$\text{L/S} = \int_0^\tau \exp(-\tilde{\kappa}_1[\tau - s]) \left(\int_s^\tau \exp(-\tilde{\kappa}_2[u - s]) du \right) ds \tag{181c}$$

$$\text{S/L} = \int_0^\tau \exp(-\tilde{\kappa}_2[\tau - s]) \left(\int_s^\tau \exp(-\tilde{\kappa}_1[u - s]) du \right) ds \tag{181d}$$

I calculate the expression for S/L explicitly, and the other results follow from appropriate substitutions. Hence:

$$\begin{aligned}
\int_s^\tau \exp(-\tilde{\kappa}_2[u - s]) du &= -\frac{1}{\tilde{\kappa}_1} \exp(-\tilde{\kappa}_1[u - s]) \Big|_s^\tau \\
&= \frac{1 - \exp(-\tilde{\kappa}_1[\tau - s])}{\tilde{\kappa}_1}
\end{aligned} \tag{182}$$

$$\begin{aligned}
\text{S/L} &= \int_0^\tau \exp(-\tilde{\kappa}_2[\tau-s]) \cdot \frac{1 - \exp(-\tilde{\kappa}_1[\tau-s])}{\tilde{\kappa}_1} ds \\
&= \int_0^\tau \frac{1}{\tilde{\kappa}_1} \exp(-\tilde{\kappa}_2[\tau-s]) - \frac{1}{\tilde{\kappa}_1} \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2][\tau-s]) ds \\
&= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} \exp(-\tilde{\kappa}_2[\tau-s]) - \frac{1}{\tilde{\kappa}_1 [\tilde{\kappa}_1 + \tilde{\kappa}_2]} \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2][\tau-s]) \Big|_0^\tau \\
&= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} [1 - \exp(-\tilde{\kappa}_2\tau)] - \frac{1}{\tilde{\kappa}_1 [\tilde{\kappa}_1 + \tilde{\kappa}_2]} [1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2]\tau)] \quad (183)
\end{aligned}$$

The result for L/L is obtained by setting $\tilde{\kappa}_2 = \tilde{\kappa}_1$ in equation 183, i.e.:

$$\begin{aligned}
\text{L/L} &= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_1} [1 - \exp(-\tilde{\kappa}_1\tau)] - \frac{1}{\tilde{\kappa}_1 [\tilde{\kappa}_1 + \tilde{\kappa}_1]} [1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_1]\tau)] \\
&= \frac{1}{\tilde{\kappa}_1^2} [1 - \exp(-\tilde{\kappa}_1\tau)] - \frac{1}{2\tilde{\kappa}_1^2} [1 - \exp(-2\tilde{\kappa}_1\tau)] \\
&= \frac{1}{2\tilde{\kappa}_1^2} [2 - 2\exp(-\tilde{\kappa}_1\tau) - 1 + \exp(-2\tilde{\kappa}_1\tau)] \\
&= \frac{1}{2\tilde{\kappa}_1^2} [1 - 2\exp(-\tilde{\kappa}_1\tau) + \exp(-2\tilde{\kappa}_1\tau)] \\
&= \frac{1}{2\tilde{\kappa}_1^2} [1 - \exp(-\tilde{\kappa}_1\tau)]^2 \\
&= \frac{1}{2} [G(\tilde{\kappa}_1, \tau)]^2 \quad (184)
\end{aligned}$$

The result for S/S is obtained by setting $\tilde{\kappa}_1 = \tilde{\kappa}_2$ in equation 184, i.e.:

$$\text{S/S} = \frac{1}{2} [G(\tilde{\kappa}_2, \tau)]^2 \quad (185)$$

The result for L/S+S/L is obtained by switching $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ in equation 183 and

adding the result for S/L, i.e.:

$$\begin{aligned}
L/S + S/L &= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} [1 - \exp(-\tilde{\kappa}_2 \tau)] - \frac{1}{\tilde{\kappa}_2 [\tilde{\kappa}_1 + \tilde{\kappa}_2]} [1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] \tau)] \\
&\quad + \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} [1 - \exp(-\tilde{\kappa}_1 \tau)] - \frac{1}{\tilde{\kappa}_1 [\tilde{\kappa}_1 + \tilde{\kappa}_2]} [1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] \tau)] \\
&= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} [2 - \exp(-\tilde{\kappa}_1 \tau) - \exp(-\tilde{\kappa}_2 \tau)] \\
&\quad - \frac{\tilde{\kappa}_1}{\tilde{\kappa}_1 \tilde{\kappa}_2 [\tilde{\kappa}_1 + \tilde{\kappa}_2]} [1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] \tau)] \\
&\quad - \frac{\tilde{\kappa}_2}{\tilde{\kappa}_1 \tilde{\kappa}_2 [\tilde{\kappa}_1 + \tilde{\kappa}_2]} [1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] \tau)] \\
&= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} [2 - \exp(-\tilde{\kappa}_1 \tau) - \exp(-\tilde{\kappa}_2 \tau)] \\
&\quad - \frac{[\tilde{\kappa}_1 + \tilde{\kappa}_2]}{\tilde{\kappa}_1 \tilde{\kappa}_2 [\tilde{\kappa}_1 + \tilde{\kappa}_2]} [1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] \tau)] \\
&= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} \{1 - \exp(-\tilde{\kappa}_1 \tau) + 1 - \exp(-\tilde{\kappa}_2 \tau) - 1 - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] \tau)\} \\
&= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} \{1 - \exp(-\tilde{\kappa}_1 \tau) - \exp(-\tilde{\kappa}_2 \tau) - \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] \tau)\} \\
&= \frac{1}{\tilde{\kappa}_1 \tilde{\kappa}_2} [1 - \exp(-\tilde{\kappa}_1 \tau)] [1 - \exp(-\tilde{\kappa}_2 \tau)] \\
&= G(\tilde{\kappa}_1, \tau) G(\tilde{\kappa}_2, \tau)
\end{aligned}$$

The final result for the volatility effect is therefore:

$$\text{VE}(\tau) = \sigma_1^2 \cdot \frac{1}{2} [G(\tilde{\kappa}_1, \tau)]^2 + \sigma_2^2 \cdot \frac{1}{2} [G(\tilde{\kappa}_2, \tau)]^2 + \rho_{12} \sigma_1 \sigma_2 \cdot G(\tilde{\kappa}_1, \tau) G(\tilde{\kappa}_2, \tau) \quad (186)$$

The square of the annualized option volatility is calculated from the forward rate factor loadings and volatility matrix as follows:

$$\begin{aligned}
[\omega(\tau)]^2 &= \int_0^\tau [\mathfrak{f}(b_0, \tilde{\kappa}, s)]' \sigma' \sigma [\mathfrak{f}(b_0, \tilde{\kappa}, s)] \, ds \\
&= \int_0^\tau [\exp(-\tilde{\kappa}_1 s), \exp(-\tilde{\kappa}_2 s)] \begin{bmatrix} \sigma_1^2 & \rho_{12} \sigma_1 \sigma_2 \\ \rho_{12} \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} \exp(-\tilde{\kappa}_1 s) \\ \exp(-\tilde{\kappa}_2 s) \end{bmatrix} \, ds \\
&= \int_0^\tau \{ \sigma_1^2 \cdot \exp(-2\tilde{\kappa}_1 s) + \sigma_2^2 \cdot \exp(-2\tilde{\kappa}_2 s) + 2\rho_{12} \sigma_1 \sigma_2 \cdot \exp(-[\tilde{\kappa}_1 + \tilde{\kappa}_2] s) \} \, ds \\
&= \sigma_1^2 \cdot G(2\tilde{\kappa}_1, \tau) + \sigma_2^2 \cdot G(2\tilde{\kappa}_2, \tau) + 2\rho_{12} \sigma_1 \sigma_2 G(\tilde{\kappa}_1 + \tilde{\kappa}_2, \tau) \quad (187)
\end{aligned}$$

Note that the results above may readily be extended to any shadow-GATSM(2) with a diagonal $\tilde{\kappa}$. Furthermore, shadow-GATSMs with a non-diagonal matrix $\tilde{\kappa}$ may be diagonalized using an eigensystem decomposition and an affine invariant transformation (see Dai and Singleton (2002) footnote 9 on p. 438, or Singleton (2006) p. 319-20 for a more general description of affine invariant transformations). Therefore, the results presented here may be applied to represent any GATSM with distinct eigenvalues for $\tilde{\kappa}$.

G Shadow-AFNSM(2)

The shadow-AFNSM(2) could be obtained directly as in the previous section by setting $a_0 = 0$ and $\tilde{\kappa} = \text{diag}[0, \phi]$, which would produce the first two forward rate factor loadings for the AFNSM(2). Specifically:

$$\begin{aligned}
 [\mathbb{1}(b_0, \tilde{\kappa}, \tau)]' &= b'_0 \exp(-\tilde{\kappa}\tau) \\
 &= [1, 1] \exp\left(-\begin{bmatrix} 0 & 0 \\ 0 & \phi \end{bmatrix} \tau\right) \\
 &= [1, 1] \begin{bmatrix} 1 & 0 \\ 0 & \exp(-\phi\tau) \end{bmatrix} \\
 &= [1, \exp(-\phi\tau)]
 \end{aligned} \tag{188}$$

where 1 and $\exp(-\phi\tau)$ are respectively the Level and Slope factor loadings for the AFNSM(2) in its forward rate form. Those factor loadings could then be used to directly calculate the volatility effect and annualized option volatility as in the previous section.

Alternatively, the results may be obtained more readily by taking the limit of $\tilde{\kappa}_1 \rightarrow 0$ for the results in the previous section and setting $a_0 = 0$. Hence, for the $\tilde{\mathbb{E}}_t[r(t + \tau)]$ expression $\lim_{\tilde{\kappa}_1 \rightarrow 0} \exp(-\tilde{\kappa}_1\tau) = 1$, and substituting $\phi = \tilde{\kappa}_2$ for notational convenience gives:

$$\tilde{\mathbb{E}}_t[r(t + \tau)] = x_1(t) + x_2(t) \cdot \exp(-\phi\tau) \tag{189}$$

Regarding the volatility effect, $\lim_{\tilde{\kappa}_1 \rightarrow 0} G(\tilde{\kappa}_1, \tau)$ is:

$$\begin{aligned}
 \lim_{\tilde{\kappa}_1 \rightarrow 0} G(\tilde{\kappa}_1, \tau) &= \lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ \frac{1}{\tilde{\kappa}_1} [1 - \exp(-\tilde{\kappa}_1\tau)] \right\} \\
 \langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ \frac{d}{d\tilde{\kappa}_1} [1 - \exp(-\tilde{\kappa}_1\tau)] \right\}}{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ \frac{d}{d\tilde{\kappa}_1} \tilde{\kappa}_1 \right\}} \\
 &= \frac{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ \tau \exp(-\tilde{\kappa}_1\tau) \right\}}{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ 1 \right\}} \\
 &= \tau
 \end{aligned} \tag{190}$$

and therefore, with $\phi = \tilde{\kappa}_2$, equation 186 becomes:

$$\text{VE}(\tau) = \sigma_1^2 \cdot \frac{1}{2} \tau^2 + \sigma_2^2 \cdot \frac{1}{2} [G(\phi, \tau)]^2 + \rho_{12} \sigma_1 \sigma_2 \cdot \tau G(\phi, \tau) \tag{191}$$

The square of the annualized option volatility may be obtained using the same evaluation of $\lim_{\tilde{\kappa}_1 \rightarrow 0} G(\tilde{\kappa}_1, \tau)$, and that for $\lim_{\tilde{\kappa}_1 \rightarrow 0} G(2\tilde{\kappa}_1, \tau)$, i.e.:

$$\begin{aligned}
 \lim_{\tilde{\kappa}_1 \rightarrow 0} G(2\tilde{\kappa}_1, \tau) &= \lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ \frac{1}{2\tilde{\kappa}_1} [1 - \exp(-2\tilde{\kappa}_1\tau)] \right\} \\
 \langle \text{L'Hopital's rule} \rangle &= \frac{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ \frac{d}{d\tilde{\kappa}_1} [1 - \exp(-2\tilde{\kappa}_1\tau)] \right\}}{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ \frac{d}{d\tilde{\kappa}_1} [2\tilde{\kappa}_1] \right\}} \\
 &= \frac{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ 2\tau \exp(-2\tilde{\kappa}_1\tau) \right\}}{\lim_{\tilde{\kappa}_1 \rightarrow 0} \left\{ 2 \right\}} \\
 &= \tau
 \end{aligned} \tag{192}$$

and so, with $\phi = \tilde{\kappa}_2$, equation 187 becomes:

$$[\omega(\tau)]^2 = \sigma_1^2 \cdot \tau + \sigma_2^2 \cdot G(2\phi, \tau) + 2\rho_{12}\sigma_1\sigma_2G(\phi, \tau) \quad (193)$$

H Shadow-AFNSM(3)

With $b_0 = [1, 1, 0]$ and:

$$\tilde{\kappa} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \phi & -\phi \\ 0 & 0 & \phi \end{bmatrix} \quad (194)$$

the shadow-AFNSM(3) forward rate factor loading vector function $[\mathfrak{f}(b_0, \tilde{\kappa}, \tau)]$ becomes:

$$\begin{aligned} [\mathfrak{f}(b_0, \tilde{\kappa}, \tau)] &= b'_0 \exp(-\tilde{\kappa}\tau) \\ &= [1, 1, 0] \exp\left(-\begin{bmatrix} 0 & 0 & 0 \\ 0 & \phi & -\phi \\ 0 & 0 & \phi \end{bmatrix} \tau\right) \end{aligned} \quad (195)$$

Following Moler and Van Loan (2003) p. 24, the matrix exponential may be evaluated via a decomposition into Jordan canonical form. That is, if A has the the Jordan decomposition:

$$A = X[\mathbf{J}_1 \oplus \dots \oplus \mathbf{J}_k]X^{-1} \quad (196)$$

where \oplus is the Kronecker sum (see, for example Higam (2008) p. 331, or below for a practical example of its application) then the matrix exponential of A is:

$$\exp(\tau A) = X[\exp(\tau \mathbf{J}_1) \oplus \dots \oplus \exp(\tau \mathbf{J}_k)]X^{-1} \quad (197)$$

The required Jordan decomposition for the AFNSM(3) $\tilde{\kappa}$ is:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & \phi & -\phi \\ 0 & 0 & \phi \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{1}{\phi} & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \phi & 0 \\ 0 & 1 & \phi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -\phi \\ 0 & 1 & 0 \end{bmatrix} \quad (198)$$

which has two Jordan blocks, i.e. $\mathbf{J}_1 = [0]$, and:

$$\mathbf{J}_2 = \begin{bmatrix} \phi & 0 \\ 1 & \phi \end{bmatrix} \quad (199)$$

The matrix exponentials of the two blocks are respectively $\exp(-\mathbf{J}_1\tau) = 1$, and:

$$\begin{aligned} \exp(-\mathbf{J}_2\tau) &= \exp(-\phi\tau) \begin{bmatrix} 1 & 0 \\ \tau & 1 \end{bmatrix} \\ &= \begin{bmatrix} \exp(-\phi\tau) & 0 \\ \tau \exp(-\phi\tau) & \exp(-\phi\tau) \end{bmatrix} \end{aligned} \quad (200)$$

so therefore:

$$\begin{aligned}
\exp(-\tilde{\kappa}\tau) &= X \left(1 \oplus \begin{bmatrix} \exp(-\phi\tau) & 0 \\ \tau \exp(-\phi\tau) & \exp(-\phi\tau) \end{bmatrix} \right) X^{-1} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{1}{\phi} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \exp(-\phi\tau) & 0 \\ 0 & \tau \exp(-\phi\tau) & \exp(-\phi\tau) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -\phi \\ 0 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \exp(-\phi\tau) & -\phi\tau \exp(-\phi\tau) \\ 0 & 0 & \exp(-\phi\tau) \end{bmatrix} \tag{201}
\end{aligned}$$

The final expression for $[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]'$ is then:

$$\begin{aligned}
[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)] &= b'_0 \exp(-\tilde{\kappa}\tau) \\
&= [1, 1, 0] \begin{bmatrix} 1 & 0 & 0 \\ 0 & \exp(-\phi\tau) & -\phi\tau \exp(-\phi\tau) \\ 0 & 0 & \exp(-\phi\tau) \end{bmatrix} \\
&= [1, \exp(-\phi\tau), -\phi\tau \exp(-\phi\tau)] \tag{202}
\end{aligned}$$

where 1, $\exp(-\phi\tau)$, and $\phi\tau \exp(-\phi\tau)$ are respectively the Level, Slope, and Bow (or Curvature) factor loadings for the AFNSM(3) in its forward rate form.

Setting $a_0 = 0$, the expected path of the shadow short rate is:

$$\begin{aligned}
\mathbb{E}_t[r(t+\tau)] &= b'_0 \exp(-\tilde{\kappa}\tau) x(t) \\
&= [1, \exp(-\phi\tau), -\phi\tau \exp(-\phi\tau)] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \\
&= x_1 + x_2 \cdot \exp(-\phi\tau) - x_3 \cdot \phi\tau \exp(-\phi\tau) \tag{203}
\end{aligned}$$

Note that could re-define the Bow loading to enter positively in $\mathbb{E}_t[r(t+\tau)]$, as in Christensen and Rudebusch (2013), but that choice would need to be carried through consistently to the volatility effect and annualized option volatility calculations below.

With the volatility matrix:

$$\sigma' = \begin{bmatrix} \sigma_1 & 0 & 0 \\ \sigma_2 \rho_{12} & \sigma_2 \sqrt{1 - \rho_{12}^2} & 0 \\ \sigma_3 \rho_{13} & \sigma_3 \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1 - \rho_{12}^2}} & \sigma_3 \sqrt{1 - \rho_{13}^2 - \frac{(\rho_{23} - \rho_{12}\rho_{13})^2}{1 - \rho_{12}^2}} \end{bmatrix} \tag{204}$$

the forward rate volatility loadings are defined as $\sigma(\tau) = \sigma[\mathfrak{h}(b_0, \tilde{\kappa}, \tau)]$ and so the

volatility effect is:

$$\begin{aligned}
\text{VE}(\tau) &= \int_0^\tau [\mathfrak{h}(b_0, \tilde{\kappa}, \tau - s)]' \sigma' \left[\int_s^\tau \sigma [\mathfrak{h}(b_0, \tilde{\kappa}, s - u)] du \right] ds \\
&= \int_0^\tau \left(\begin{bmatrix} \text{L} \\ \text{S} \\ \text{B} \end{bmatrix}' \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix} \int_s^\tau \begin{bmatrix} \text{L} \\ \text{S} \\ \text{B} \end{bmatrix} du \right) ds \\
&= \begin{bmatrix} \sigma_1^2 \cdot \text{L/L} & +\rho_{12}\sigma_1\sigma_2 \cdot \text{L/S} & +\rho_{13}\sigma_1\sigma_3 \cdot \text{L/B} \\ +\rho_{12}\sigma_1\sigma_2 \cdot \text{S/L} & +\sigma_2^2 \cdot \text{S/S} & +\rho_{23}\sigma_2\sigma_3 \cdot \text{S/B} \\ +\rho_{13}\sigma_1\sigma_3 \cdot \text{B/L} & +\rho_{23}\sigma_2\sigma_3 \cdot \text{B/S} & +\sigma_3^2 \cdot \text{B/B} \end{bmatrix} \\
&= \sigma_1^2 \cdot \text{L/L} + \sigma_2^2 \cdot \text{S/S} + \rho_{12}\sigma_1\sigma_2 \cdot (\text{L/S} + \text{S/L}) \\
&\quad + \sigma_3^2 \cdot \text{B/B} - \rho_{13}\sigma_1\sigma_3 \cdot (\text{L/B} + \text{B/L}) \\
&\quad - \rho_{23}\sigma_2\sigma_3 \cdot (\text{S/B} + \text{B/S}) \tag{205}
\end{aligned}$$

where L, S, and B are abbreviated notation for the functions $[s - u]$, $\exp(-\phi[s - u])$ and $\phi[s - u] \exp(-\phi[s - u])$ and their $[\tau - s]$ counterparts in the second integral, and the combinations are analogous to the individual integrals for the GATSM(2) specified in appendix C.2.

The components $\sigma_1^2 \cdot \text{L/L}$, $\sigma_2^2 \cdot \text{S/S}$, $\rho_{12}\sigma_1\sigma_2 \cdot (\text{L/S} + \text{S/L})$ are already available for the AFSNM(2), and the remaining components are calculated as follows. For the B/B component:

$$\begin{aligned}
\text{B/B} &= \int_0^\tau \phi[\tau - s] \exp(-\phi[\tau - s]) \left(\int_s^\tau \phi[u - s] \exp(-\phi[u - s]) du \right) ds \\
&= \frac{1}{2} \left(\frac{1}{\phi} [1 - \exp(-\phi\tau)] - \tau \exp(-\phi\tau) \right)^2 \\
&= \frac{1}{2} [G(\phi, \tau) - \tau \exp(-\phi\tau)]^2 \\
&= \frac{1}{2} [F(\phi, \tau)]^2 \tag{206}
\end{aligned}$$

where:

$$F(\phi, \tau) = G(\phi, \tau) - \tau \exp(-\phi\tau) \tag{207}$$

For the L/B + B/L component:

$$\begin{aligned}
\text{L/B} + \text{B/L} &= \int_0^\tau 1 \left(\int_s^\tau \phi[u - s] \exp(-\phi[u - s]) du \right) ds \\
&\quad + \int_0^\tau \phi[\tau - s] \exp(-\phi[\tau - s]) \left(\int_s^\tau 1 du \right) ds \\
&= \tau \frac{1 - \exp(-\phi\tau)}{\phi} - \tau^2 \exp(-\phi\tau) \\
&= \tau G(\phi, \tau) - \tau^2 \exp(-\phi\tau) \\
&= \tau [G(\phi, \tau) - \tau \exp(-\phi\tau)] \\
&= \tau F(\phi, \tau) \tag{208}
\end{aligned}$$

For the S/B + B/S component:

$$\begin{aligned}
\text{S/B} + \text{B/S} &= \int_0^\tau \exp(-\phi[\tau-s]) \left(\int_s^\tau \phi[u-s] \exp(-\phi[u-s]) du \right) ds \\
&\quad + \int_0^\tau \phi[\tau-s] \exp(-\phi[\tau-s]) \left(\int_s^\tau \exp(-\phi[u-s]) du \right) ds \\
&= \left(\frac{1 - \exp(-\phi\tau)}{\phi} \right)^2 - \tau \exp(-\phi\tau) \frac{1 - \exp(-\phi\tau)}{\phi} \\
&= [G(\phi, \tau)]^2 - \tau \exp(-\phi\tau) G(\phi, \tau) \\
&= G(\phi, \tau) [G(\phi, \tau) - \tau \exp(-\phi\tau)] \\
&= G(\phi, \tau) F(\phi, \tau) \tag{209}
\end{aligned}$$

Therefore, the volatility effect for the AFNSM(3) is:

$$\begin{aligned}
\text{VE}(\tau) &= \sigma_1^2 \cdot \tau^2 + \sigma_2^2 \cdot \frac{1}{2} [G(\phi, \tau)]^2 + \sigma_3^2 \cdot \frac{1}{2} [F(\phi, \tau)]^2 \\
&\quad + \rho_{12} \sigma_1 \sigma_2 \cdot \tau G(\phi, \tau) - \rho_{13} \sigma_1 \sigma_3 \cdot \tau F(\phi, \tau) \\
&\quad - \rho_{23} \sigma_2 \sigma_3 \cdot G(\phi, \tau) F(\phi, \tau) \tag{210}
\end{aligned}$$

The square of the annualized option volatility is calculated from the forward rate factor loadings and the volatility matrix as follows:

$$\begin{aligned}
[\omega(\tau)]^2 &= \int_0^\tau [\mathfrak{h}(b_0, \tilde{\kappa}, s)]' \sigma' \sigma [\mathfrak{h}(b_0, \tilde{\kappa}, s)] ds \\
&= \int_0^\tau \begin{bmatrix} \text{L} \\ \text{S} \\ \text{B} \end{bmatrix}' \begin{bmatrix} \sigma_1^2 & \rho_{12} \sigma_1 \sigma_2 & \rho_{13} \sigma_1 \sigma_3 \\ \rho_{12} \sigma_1 \sigma_2 & \sigma_2^2 & \rho_{23} \sigma_2 \sigma_3 \\ \rho_{13} \sigma_1 \sigma_3 & \rho_{23} \sigma_2 \sigma_3 & \sigma_3^2 \end{bmatrix} \begin{bmatrix} \text{L} \\ \text{S} \\ \text{B} \end{bmatrix} ds \\
&= \int_0^\tau [\sigma_1^2 + \sigma_2^2 \cdot \exp(-2\phi u) + 2\rho_{12} \sigma_1 \sigma_2 \cdot \exp(-\phi u)] du \tag{211a}
\end{aligned}$$

$$+ \int_0^\tau \sigma_3^2 \cdot (\phi u)^2 \exp(-2\phi u) du \tag{211b}$$

$$- \int_0^\tau 2\rho_{13} \sigma_1 \sigma_3 \cdot \phi u \exp(-\phi u) du \tag{211c}$$

$$- \int_0^\tau 2\rho_{23} \sigma_2 \sigma_3 \cdot \phi u \exp(-2\phi u) du \tag{211d}$$

The result for equation 211a is already available from appendix C.3, and the remaining components for equation 211 are calculated as follows. For 211b:

$$\begin{aligned}
\int_0^\tau (\phi u)^2 \exp(-2\phi u) du &= \frac{1}{2} [G(2\phi, \tau) - \tau \exp(-2\phi\tau) - \phi\tau^2 \exp(-2\phi\tau)] \\
&= \frac{1}{2} [F(2\phi, \tau) - \phi\tau^2 \exp(-2\phi\tau)] \tag{212}
\end{aligned}$$

where:

$$F(2\phi, \tau) = G(2\phi, \tau) - \tau \exp(-2\phi\tau) \tag{213}$$

for 211c:

$$\begin{aligned}\int_0^\tau \phi u \exp(-2\phi u) du &= G(\phi, \tau) - \tau \exp(-\phi\tau) \\ &= F(\phi, \tau)\end{aligned}\quad (214)$$

and for 211d:

$$\begin{aligned}\int_0^\tau \phi u \exp(-\phi u) du &= G(2\phi, \tau) - \tau \exp(-2\phi\tau) \\ &= F(2\phi, \tau)\end{aligned}\quad (215)$$

Therefore, the final expression for $[\omega(\tau)]^2$ in the CAB-AFNSM(3) is:

$$\begin{aligned}[\omega(\tau)]^2 &= \sigma_1^2 \cdot \tau + \sigma_2^2 \cdot G(2\phi, \tau) + \sigma_3^2 \cdot \frac{1}{2} [F(2\phi, \tau) - \phi\tau^2 \exp(-2\phi\tau)] \\ &\quad + 2\rho_{12}\sigma_1\sigma_2 G(\phi, \tau) - 2\rho_{13}\sigma_1\sigma_3 F(\phi, \tau) - \rho_{23}\sigma_2\sigma_3 F(2\phi, \tau)\end{aligned}\quad (216)$$

I Simplified GATSM volatility effect expressions

This appendix is not referenced anywhere in the article, but I have included it to show that the shadow-GATSM volatility effect calculations for the models in this article may be simplified considerably, which is an observation that may prove useful for future CAB-GATSM or GATSM work. Specifically, it turns out that the volatility effects for the models in sections 5.2, 5.3, and 5.4 may all be obtained as the inner product of a single integral rather than a double integral. That is, defining the volatility effect as:

$$\text{VE}(\tau) = \frac{1}{2} [\mathbf{L}(b_0, \tilde{\kappa}, \tau)]' \sigma' \sigma [\mathbf{L}(b_0, \tilde{\kappa}, \tau)] \quad (217)$$

where:

$$[\mathbf{L}(b_0, \tilde{\kappa}, \tau)] = \left[\int_0^\tau [\mathbf{l}(b_0, \tilde{\kappa}, s)] ds \right] \quad (218)$$

produces the volatility effect expressions for the three shadow-GATSMs specified.

For the shadow-GATSM(2):

$$\begin{aligned}[\mathbf{L}(b_0, \tilde{\kappa}, \tau)]' &= \left[\int_0^\tau [\exp(-\tilde{\kappa}_1 s), \exp(-\tilde{\kappa}_2 s)]' ds \right] \\ &= [G(\tilde{\kappa}_1, \tau), G(\tilde{\kappa}_2, \tau)]'\end{aligned}\quad (219)$$

and:

$$\begin{aligned}\text{VE}(\tau) &= \frac{1}{2} [\mathbf{L}(b_0, \tilde{\kappa}, \tau)]' \sigma' \sigma [\mathbf{L}(b_0, \tilde{\kappa}, \tau)] \\ &= \frac{1}{2} [G(\tilde{\kappa}_1, \tau), G(\tilde{\kappa}_2, \tau)] \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} G(\tilde{\kappa}_1, \tau) \\ G(\tilde{\kappa}_2, \tau) \end{bmatrix} \\ &= \sigma_1^2 \cdot \frac{1}{2} [G(\tilde{\kappa}_1, \tau)]^2 + \sigma_2^2 \cdot \frac{1}{2} [G(\tilde{\kappa}_2, \tau)]^2 + \rho_{12}\sigma_1\sigma_2 \cdot G(\tilde{\kappa}_1, \tau) G(\tilde{\kappa}_2, \tau)\end{aligned}\quad (220)$$

For the shadow-AFNSM(2):

$$\begin{aligned} [\mathbf{L}(b_0, \tilde{\kappa}, \tau)]' &= \left[\int_0^\tau [1, \exp(-\phi s)]' ds \right] \\ &= [\tau, G(\phi, \tau)]' \end{aligned} \quad (221)$$

and:

$$\begin{aligned} \text{VE}(\tau) &= \frac{1}{2} [1, G(\phi, \tau)] \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 \\ G(\phi, \tau) \end{bmatrix} \\ &= \sigma_1^2 \cdot \frac{1}{2} \tau^2 + \sigma_2^2 \cdot \frac{1}{2} [G(\phi, \tau)]^2 + \rho_{12}\sigma_1\sigma_2 \cdot \tau G(\phi, \tau) \end{aligned} \quad (222)$$

For the shadow-AFNSM(3):

$$\begin{aligned} [\mathbf{L}(b_0, \tilde{\kappa}, \tau)]' &= \left[\int_0^\tau [1, \exp(-\phi s), \phi s \exp(-\phi s)]' ds \right] \\ &= \left[\tau, G(\phi, \tau), \frac{1}{\phi} [1 - \exp(-\phi\tau)] - \tau \exp(-\phi\tau) \right]' \\ &= [\tau, G(\phi, \tau), F(\phi, \tau)]' \end{aligned} \quad (223)$$

and:

$$\begin{aligned} \text{VE}(\tau) &= \frac{1}{2} [1, G(\phi, \tau), F(\phi, \tau)] \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_1\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_1\sigma_2 & \sigma_3^2 \end{bmatrix} \begin{bmatrix} 1 \\ G(\phi, \tau) \\ F(\phi, \tau) \end{bmatrix} \\ &= \sigma_1^2 \cdot \frac{1}{2} \tau^2 + \sigma_2^2 \cdot \frac{1}{2} [G(\phi, \tau)]^2 + \sigma_3^2 \cdot \frac{1}{2} [F(\phi, \tau)]^2 \\ &\quad + \rho_{12}\sigma_1\sigma_2 \cdot \tau G(\phi, \tau) + \rho_{13}\sigma_1\sigma_3 \cdot \tau F(\phi, \tau) \\ &\quad + \rho_{23}\sigma_2\sigma_3 \cdot G(\phi, \tau) F(\phi, \tau) \end{aligned} \quad (224)$$

I conjecture that the result holds generally for all GATSMs, and it should be straightforward to establish that proof. However, I haven't yet looked at doing so.