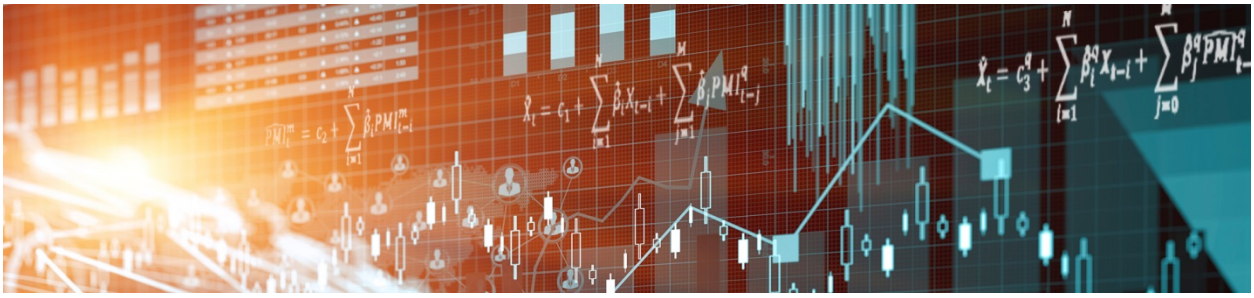
BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2015-46

Tractable Term Structure Models



by Bruno Feunou, Jean-Sébastien Fontaine, Anh Le and Christian Lundblad

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2015-46

December 2015

Last updated July 2021

Tractable Term Structure Models

by

Bruno Feunou,[1] Jean-Sébastien Fontaine,[1] Anh Le[2] and Christian Lundblad[3]

[1]Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada
Corresponding author: jsfontaine@bankofcanada.ca

[2]Smeal College of Business
Pennsylvania State University

[3]Kenan–Flagler Business School
University of North Carolina at Chapel Hill

ISSN 1701-9397

Acknowledgements

We thank for comments and suggestions Peter Christoffersen, Greg Duffee, Antonio Diez de los Rios, Peter Feldhutter, Jing-zhi Huang, Kris Jacobs, Richard Luger, and seminar and conference participants at the Banque de France, the Federal Reserve Bank of New York, McGill University, the National Bank of Belgium, Penn State Smeal College of Business, Tsinghua University PBC School of Finance, and the 2014 Tremblant Risk Conference, 2015 CIRANO finance workshop, 2015 Bank of Canada–FRBSB Fixed Income Conference, 2017 MFA meetings. We thank Alison Arnot and Meredith Fraser-Ohman for editorial support.

Abstract

We introduce a new framework that facilitates term structure modeling with both positive interest rates and flexible time-series dynamics but that is also tractable, meaning amenable to quick and robust estimation. Using both simulations and U.S. historical data, we compare our approach with benchmark Gaussian and stochastic volatility models as well as a shadow rate model that enforces positive interest rates. Our approach, which remains arbitrarily close to arbitrage-free, offers a more accurate characterization of bond Sharpe ratios due to a better fit of the volatility dynamics and a more efficient estimation of the return dynamics. Further, standard shadow rate and stochastic volatility models exhibit important restrictions that are largely absent in our approach.

JEL classification: G12
Bank classification: Asset pricing; Interest rates; Transmission of monetary policy; Uncertainty and monetary policy; International topics; International financial markets

Résumé

Nous présentons un nouveau cadre qui facilite la modélisation de la structure par terme des taux d'intérêt en tenant compte à la fois des taux d'intérêt positifs et des dynamiques flexibles associées aux séries chronologiques. Ce cadre se distingue par sa maniabilité, à savoir qu'il produit rapidement des estimations robustes. En utilisant des simulations et les données historiques disponibles pour les États-Unis, nous comparons notre approche à celle de modèles de référence gaussiens ou à volatilité stochastique, mais aussi d'un modèle similaire à celui de Black (1995), qui impose une condition de taux d'intérêt positifs. Notre approche coïncide dans les grandes lignes avec celle de modèles imposant l'absence d'opportunités d'arbitrage, mais autorise une meilleure représentation des ratios de Sharpe obligataires, du fait que la dynamique de la volatilité est plus fidèlement reflétée et que la dynamique des rendements est estimée plus efficacement. Il faut également souligner que tous les modèles de référence que nous considérons comportent d'importantes restrictions qui sont essentiellement absentes de notre approche.

Classification JEL : G12
Classification de la Banque : Évaluation des actifs; Taux d'intérêt; Transmission de la politique monétaire; Incertitude et politique monétaire; Questions internationales; Marchés financiers internationaux

1 Introduction

As interest rates remain at or below zero in the face of persistent unconventional monetary policy in many industrialized countries, the need for a tractable term structure model — one that is amenable to quick and robust estimation — that can accommodate the near lower bound behavior of interest rates is stronger than ever before. Despite its practical importance even after years at the lower bound, characterizing bond risks and returns remains vexing. Unfortunately, existing bond pricing models that incorporate bounded interest rates (starting with Black, 1995) suffer from intractable bond pricing, inflexible volatility dynamics, or both. Our contribution is to offer a framework to build potentially non-linear models that are tractable yet arbitrarily close to arbitrage-free.

Subjecting this framework to both realistic data simulations and U.S. historical data, we find significant improvements relative to benchmark models along several important dimensions. In a simulation environment, we show that our approach offers more accurate characterizations of bond Sharpe ratios. By combining information on risks and returns, Sharpe ratios are an important statistic to differentiate models on economic grounds. Second, in U.S. historical data, we find that our tractable models offer yield and return forecasts with the same or better accuracy than other benchmark models and improve yield volatility forecasts. These results occur both away from and near the lower bound, as well as in both in-sample and out-of-sample settings.

The elaborate simulation and out-of-sample evidence also provides valuable empirical insights that are relevant for researchers, policy-makers, and portfolio managers alike. First, researchers have long struggled to obtain an accurate decomposition of risk premia in terms of prices and quantities of risk. It is well known that, even in normal times, many existing models face a tension in simultaneously capturing both components accurately; this is only more complicated during zero lower bound (ZLB) episodes. In its ability to capture volatility dynamics without compromising the estimation of risk premia, our framework offers a tractable avenue to overcoming this tension. Second, policy-makers are also interested in the ability of a model to provide reliable forecasts of these components over multiple horizons to better inform, for example, monetary policy and systemic risk analysis.

Finally, portfolio managers, particularly those with a large fraction of their portfolios invested in U.S. Treasuries, will find value in a model that is able to provide reliable reward-to-risk metrics such as the Sharpe ratio.[1]

[1]To provide an important example, we highlight the "official" foreign holdings of U.S. federal debt; roughly

To guarantee tractability in pricing, we directly specify a recursive construction that connects bond prices at successive maturities to a small set of state variables, consistent with the pronounced evidence that bond yields are driven by a small number of factors. The direct specification of bond prices that we employ is reminiscent of the dynamic Nelson-Siegel approach (henceforth DNS, see Diebold and Li (2006)). While the DNS is nested in our framework, we allow for a high degree of flexibility in modeling the time series dynamics of the pricing factors, including stochastic volatility, along with nearly complete freedom in modeling the short rate, including non-linear formulations that are consistent with a lower bound.

Although the pricing functions are explicitly specified, we recover prices that remain economically sensible. approach. One reason put forth in Christensen, Diebold, and Rudebusch (2011) as to why DNS models are popular is that DNS bond prices appear to be *almost* arbitrage-free even if they are not exactly free of arbitrage in a frictionless market (Bjork and Christensen, 1999; Filipovic, 1999). We pursue this intuition formally for our more general setting, showing that our broad family of models delivers bond prices that are free of dominant trading strategies in the sense of Rothschild and Stiglitz (1970) and Levy (1992). Intuitively speaking, the no-dominance property rules out almost all arbitrage opportunities, and we show that any remaining bond trading arbitrage opportunities cannot be self-financing and do not survive in the presence of non-zero short-selling costs. Overall, this analysis suggests that the practical relevance of DNS models generalizes to our broader, more flexible framework.

While our framework precludes dominant bond trading strategies under far more general assumptions, we consider an example model that facilitates comparison with existing models. Specifically, in the spirit of Black (1995) and Wu and Xia (2016), we impose a "hockey stick"-like lower bound on interest rates. Second, we permit a multivariate conditional volatility process similar in spirit to the multivariate volatility models introduced by Noureldin et al. (2011). We show how to undo the non-linearity of this example model, typically a major obstacle in estimation, and to take advantage of recent advances (see Joslin et al. (2011)) such that estimation is highly convenient, fast, and robust.

To benchmark performance, we include in our analysis three popular alternatives. First, we consider the no-arbitrage affine Gaussian term structure model that does not allow for

\$4 trillion in 2020 represents, in part, the foreign currency reserve portfolios of various central banks (U.S. Treasury Bulletin). As these reserve portfolio managers face tight risk management restrictions that often preclude credit risk or derivatives, they tend to rely on easy-to-implement term structure models to manage their high-quality government debt duration risk exposures.

a lower bound and features constant yield volatility. Second, we consider the no-arbitrage shadow rate model in Black (1995) and Wu and Xia (2016) that enforces a lower bound but also features essentially constant yield volatility away from the lower bound.[2] Third, we consider a standard no-arbitrage affine stochastic volatility model (Dai and Singleton, 2000) that does not facilitate a lower bound. Importantly, robust global estimates of all three models can be obtained in a straightforward manner.

To compare Sharpe ratio estimates across our model and these benchmarks, we set up a simulation environment where, in contrast to the case of historical data, the true conditional Sharpe ratios are known. To generate simulations, we use a data generating process that incorporates the salient features of the U.S. bond market data—such as an upward sloping average yield curve, downward sloping average yield volatility, predictable bond excess returns, time-varying volatility—that includes episodes at the lower bound toward the end of each sample. The results show that our approach offers a more accurate characterization of bond Sharpe ratios than the benchmark models.

One of the deeper insights from our results is that the Sharpe ratio estimates are more accurate because the flexible volatility dynamics in our framework readily translate into more accurate return forecasts (the numerator of the Sharpe ratio) which are due to the econometric efficiency gained by the much-improved volatility forecasts (the denominator of the Sharpe ratio). Another reason why our framework delivers more accurate Sharpe ratios is that the standard shadow rate model exhibits a pattern of biases in return and volatility forecasts that substantially aggravates the accuracy of bond and portfolio Sharpe ratio estimates. Put differently, we document that a full no-arbitrage implementation of the shadow rate model introduces a tension between returns and volatility forecasts. A related tension is well known for models relying on square-root processes (Dai and Singleton, 2000; Joslin and Le, 2021). Hence, one surprising finding is that both stochastic volatility and Black shadow rate models suffer from tensions due to the no-arbitrage restrictions binding the bond return and volatility dynamics, even if the exact nature of these tensions differs across these two models.

Moving beyond the simulation setup, we fit the same models to the U.S. historical data, in both full and out-of-sample contexts. Interestingly, we obtain pricing errors and yield forecasts that are very similar, in a statistical sense, across candidate models (both near

[2]A rich literature provides accurate and tractable approximation schemes to easily estimate shadow rate models with constant volatility. In particular, Krippner (2011), Priebsch (2013), Kim and Priebsch (2013) and Wu and Xia (2016) provided early practical approximation schemes for the model in Black (1995) and analyzed their accuracy.

and away from the ZLB). In contrast to the simulation exercise, estimates of conditional first moments may not allow for a clean separation of models in small samples, possibly due to the near unit-root nature of bond yields. In sharp contrast, we do find that our model offers a far better fit of yield volatilities, both in and out of sample (and near and away from the ZLB). Of course, including flexible volatility dynamics contributes to the improvement. For example, the benchmark no-arbitrage stochastic volatility model that we consider offers, in some sub-samples at least, fairly competitive volatility forecasts. However, as in our simulations, easing the tension between fitting yields and volatilities also plays an important role. Despite some ability to forecast volatility, the benchmark no-arbitrage stochastic volatility model fails to capture bond empirical risk premia dynamics on a real-time basis; our preferred model does not suffer this limitation.

By way of comparison, it should also be noted that Engle, Roussellet, and Siriwardane (2017) design an important extension of the dynamic Nelson-Siegel model tailored for long-term scenario generation and forecasts. Their design incorporates a lower bound, time-varying volatility and correlation, as well as regime changes that increase persistence. Our goal is different. We provide a broad framework for tractable models precluding dominant bond trading strategies across a rich set of dynamic specifications. Instead of long-horizon forecasts, our empirical application focuses on conditional Sharpe ratio estimates.

Two other important papers combine time-varying volatility with a lower bound in no-arbitrage models. Monfort et al. (2017) allow for a combination of stochastic volatility and the lower bound for interest rates within the affine no-arbitrage framework. Filipovic, Larsson, and Trolle (2017) introduce the class of linear-rational no-arbitrage models featuring several realistic features. Each of these models requires a particular combination of state dynamics and the pricing kernel to deliver analytical bond prices. Importantly, they often rely on filtering and estimation procedures, including possibly using auxiliary financial market data, that can be computationally challenging to implement in practical day-to-day applications.

Finally, a substantial body of work uses no-arbitrage shadow rate models to study bond yields and bond risk premia.[3] Without exception, this stream of literature relies on constant variances and correlations for the risk factors. However, studying the risk premium and volatility jointly has become more important given how pervasive the lower bound has become. As Bauer and Rudebusch (2016) note, the boundary introduces an asymmetry in the relationship between bond yields and bond risk factors, implying that the bond volatility

[3]Important examples include Kim and Singleton (2012), Bauer and Rudebusch (2016), Christensen and Rudebusch (2014), Krippner (2013), Wu and Xia (2016), and Priebsch (2013).

influences many important model forecasts. Recognizing this issue, Cieslak and Povala (2016) end their sample in 2010 before long-term yields, burdened by the tightness of the boundary, become less sensitive to news (Bauer and Rudebusch, 2016; Swanson and Williams, 2014). Similarly, Creal and Wu (2015) introduce time-varying macroeconomic uncertainty but end their sample in mid-2012. Both models do not prevent negative yields.

2 Tractable Dynamic Term Structure Models

This section introduces a new family of dynamic term structure models in which bond prices are specified directly, guaranteeing that bond yields are available in closed form with only minimal assumptions about the risk factor dynamics. In particular, we specify the relation between bond prices and the state vector, along with the associated state dynamics, in a way that bond yields are free of dominant trading strategies (which we will explain). Our framework admits as a special case the popular dynamic Nelson-Siegel approach of Diebold and Li (2006). However, our framework is significantly more general in its ability to accommodate realistic features of the data. In our empirical analysis, we employ a version of our proposed modeling framework that allows for the presence of a lower bound on yields as well as flexible volatility dynamics for yields.

2.1 A No-Dominance Model with a Lower Bound

Consider zero-coupon bonds with a face value of one dollar. Assumption 1 is a direct specification for the price of the n-period bond $P_n(X_t)$, where X_t is the state vector.

Assumption 1. *The n-period zero-coupon bond price $P_n(X_t)$ is given recursively by:*

$$P_n(X_t) = P_{n-1}(g(X_t)) \times \exp(-m(X_t)), \tag{1}$$

$$P_0(X_t) \equiv 1, \tag{2}$$

where $X \in \underline{X} \subseteq \mathbf{R}^N$, where $m(X) \in \underline{M} \subseteq \mathbb{R}$ is a scalar and $g(X) \in \underline{X}$.

Because zero-coupon bond prices are available in closed form, all yields and forward rates are also available. The forward rate $f_{n,t}$ applicable to a one-period loan n periods in the future is given by:

$$f_{n,t} = m(g^{\circ n}(X_t)), \tag{3}$$

where $g^{\circ k}(X_t)$ is the function $g(\cdot)$ applied k times (see Appendix A.1).

Appendix A.1 shows that this general framework guarantees the absence of dominant trading strategies (Rothschild and Stiglitz, 1970; Levy, 1992). That is, prices generated by Assumption 1 always satisfy the no-dominance (ND) condition, a concept closely related to the well-known no-arbitrage (NA) condition. Specifically, the NA condition holds if any portfolio with positive cash flows for a strictly positive probability and zero cash flows otherwise must command a strictly positive price. The weaker ND condition holds if and only if portfolios with strictly positive payoffs in all states must have positive prices. Section 2.2 provides more detail on the ND condition with respect to tractable term structure models.

The functions $m(\cdot)$ and $g(\cdot)$ in Equation (1) are the key primitives in our construction of bond prices. Applying $n = 1$ to this equation leads to the one period bond price of $P_1(X_t) = \exp(-m(X_t))$. Thus, the function $m(X_t)$ captures the one period interest rate as a function of the state vector X_t. The function $g(\cdot)$ embodies how the price of the bond is discounted back by one period. Equation (2) corresponds to our assumption that bonds are redeemed at their face value of one dollar.

Our empirical results are based on a particularly relevant example of an ND model that imposes a lower bound on interest rates in the spirit of Black (1995) and Wu and Xia (2016). For the short rate function $m(\cdot)$, we let:

$$r_t = m(X_t) = \theta\, w\left(\frac{s_t}{\theta}\right), \tag{4}$$

where $\theta > 0$ is a scalar, $s_t = \delta_0 + \delta_1' X_t$ is the shadow rate. The $w(x)$ function, closely related to the short rate function adopted by Wu and Xia (2016), is defined as:

$$w(x) = x\Phi(x) + \phi(x), \tag{5}$$

where $\Phi(x)$ and $\phi(x)$ are the cumulative probability function and the density function of a standard normal distribution, respectively.

To illustrate Equations (4)-(5), Figure 1 displays the short rate r_t for values of the shadow rate s_t between -2 percent and 5 percent where we fix $\theta = 0.0070$, producing the familiar hockey stick pattern. It is visually apparent (and it can be shown analytically) that this choice for $w(x)$ implies the following properties: (i) r_t has a lower bound, (ii) r_t increases with the shadow rate s_t, and (iii) r_t converges to the shadow rate for large values of s_t. We choose to pin down the lower bound to zero for simplicity, since the short-maturity yields in our sample do not take negative values.

As our setting is quite general, other modeling choices are available. For instance, it is possible to set $w(x) = \log(1 + \exp(x))$, which also guarantees positive interest rates. In addition, the online appendix shows other examples of ND models where yields are given by linear or linear-quadratic functions of the pricing factors.

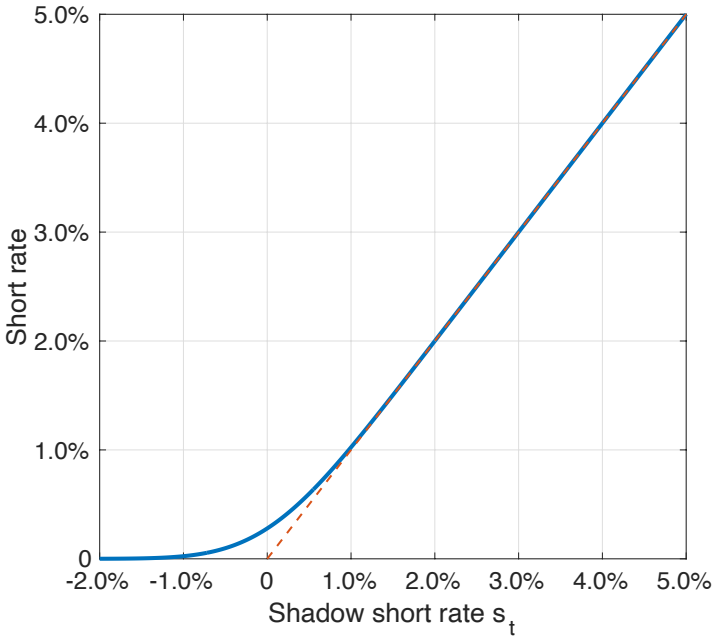


Figure 1: Hockey Stick
The hockey stick transformation from the shadow rate s_t to the short rate r_t given by $\theta w(s_t/\theta)$, where $w(x) = x\Phi(x) + \phi(x)$ and $\theta = 0.0070$. $\Phi(x)$ and $\phi(x)$ are, respectively, the cumulative probability function and the density function of a standard normal distribution.

To obtain yields and forward rates, we adopt, for simplicity, a linear choice for $g(\cdot)$:

$$g(X_t) = KX_t, \tag{6}$$

where K is a $N \times N$ matrix. While the choice of the function $w(x)$ embeds a reduced-form interpretation that s_t is the shadow short rate, Equations (4)-(6) together imply that the one-period forward rate n periods ahead is given by:

$$f_{n,t} = \theta\, w\left(\frac{\delta_0 + \delta_1' K^n X_t}{\theta}\right), \tag{7}$$

which suggests a similar interpretation linking the forward rate $f_{n,t}$ to a shadow forward

rate $s_{n,t} = \delta_0 + \delta_1' K^n X_t$. Like the short rate, the observed forward rate remains positive for all values of the shadow forward rate. Of course, the observed and shadow forward rates converge to each other for large values of the shadow rate $s_{n,t}$.

2.2 Properties

The solution for forward rates in Equation (7) is due to the recursive assumption for bond prices in Equation (1). By way of comparison, the online appendix shows that under a specific parameterization, our framework nests the Nelson-Siegel loadings. To this extent, our approach is reminiscent of the DNS framework of Diebold and Li (2006). This connection is important given the premium on tractability. While the DNS framework does not guarantee positive yields, that class of models remains popular to a large extent because it is easy to use and the models provide good forecasting performance. One reason for the popularity of DNS models is that they seem almost arbitrage-free, as put forth in Christensen, Diebold, and Rudebusch (2011).[4]

Therefore, an important question is whether more general choices of $m(\cdot)$ and $g(\cdot)$ within our framework produce bond prices that are also almost arbitrage-free and that can be trusted in practical forecasting applications. We provide an answer to this question in Appendix A.1. As mentioned above, bond prices in our framework do not offer dominant trading strategy. We show, in addition, that any remaining arbitrage opportunity our framework could permit cannot be self-financing and, in addition, that the opportunity does not survive in the presence of non-zero short-selling costs. This result holds for any choices of $m(\cdot)$ and $g(\cdot)$ and requires only mild conditions for the dynamics of the risk factors X_t. Indeed, a substantial literature has documented the costs to establish and maintain short Treasury bond positions, even for the most liquid issues (Duffie, 1996; Krishnamurthy, 2002; Vayanos and Weill, 2008; Banerjee and Graveline, 2013).

Therefore, it appears that the practical relevance of the DNS models may generalize to our broader, more flexible framework. By swapping the NA requirement for the weaker ND requirement, we hope to broaden the scope of tractable yet economically sensible term structure models. Choosing a model in the ND framework does not preclude that we recover a useful representation of the data. In addition, choosing a model in the NA framework does

[4]See Diebold and Rudebusch (2012) for an excellent survey of DNS models. Bjork and Christensen (1999) and Filipovic (1999) show theoretically that the DNS model does not preclude all arbitrage opportunities. Krippner (2013) shows it can be seen as a low-order Taylor approximation of certain no-arbitrage Gaussian affine term structure models. Coroneo et al. (2011) find that estimated parameters of DNS models are not statistically different from estimates of corresponding no-arbitrage models.

not guarantee that we recover a plausible representation of the data by all accounts. For instance, Duffee (2010) shows that estimates of fully-flexible NA models produce maximum Sharpe ratios that are astronomically high.

Beyond flexibility and tractability, the ND framework offers another potential benefit. To see this, compare the ND pricing equation (7) with what one would obtain in a standard NA setting. First, the ND intercept δ_0 is invariant to the volatility specification. The intercepts in the NA model would contain Jensen terms that vary with maturity and depend on the variance parameters. This difference resembles the distinction between the constant term in the Nelson-Siegel model and the additional Jensen term in a standard NA Gaussian model. Second, the scaling parameter θ in the ND pricing equation (7) is also invariant to the volatility specification. The scaling parameters in an NA model would vary with maturity and depend strongly on the volatility of yields. Therefore, it is plausible that the NA restriction introduces a tension between the fit of yields and estimates of the volatility parameters, as in Joslin and Le (2021) for square-root processes, that is relaxed in the ND setting.

2.3 State Dynamics

To facilitate comparison with existing models, we consider conditionally Gaussian dynamics for convenience. However, we re-emphasize that our framework precludes dominant bond trading strategies under far more general dynamics. For example, future work could easily consider additional features including a shifting endpoint in the spirit of Kozicki and Tinsley (2001); Bauer and Rudebusch (2020), a long-memory process as in Goliński and Zaffaroni (2016), or switching regimes as in Bansal et al. (2004).[5]

Every model that we consider builds on the following discrete-time vector auto-regressive VAR(1) dynamics:

$$X_{t+1} = K_0 + K_1 X_t + \Sigma_t^{1/2} \varepsilon_{t+1}, \tag{8}$$

where ε_{t+1} is i.i.d. standard normal. We also consider a general case where the conditional

[5]Additionally, the conditional mean $E_t[X_{t+1}]$ at time t may not be completely spanned by X_t. This allows for notions of unspanned risks introduced by Joslin, Priebsch, and Singleton (2014), Duffee (2011), and Feunou and Fontaine (2014). Likewise, the conditional variances $V_t[X_{t+1}]$ can be constant, as in standard Gaussian DTSMs; can depend on X_t itself, as in the $A_M(N)$ models of Dai and Singleton (2000); can depend on the history $\{X_t, X_{t-1}, \ldots\}$, in the spirit of the ARCH literature pioneered by Engle (1982); or can depend on the history of other risk factors, capturing the notion of unspanned stochastic volatility in Collin-Dufresne and Goldstein (2002), Li and Zhao (2006), and Joslin (2018).

dynamics for volatility Σ_t are given by:

$$\Sigma_t = \Sigma_0 + a\Sigma_{t-1} + b\Delta X_t \Delta X_t', \tag{9}$$

where a and b are both scalar and $\Delta X_t = X_t - X_{t-1}$, which is similar in spirit to the multivariate volatility models introduced by Noureldin, Shephard, and Sheppard (2011). We label this model SV-B_{nd}, where SV indicates time-varying volatility, B indicates a Black-type shadow rate, and the subscript $_{nd}$ indicates the no-dominance property. For comparison with the existing shadow rate models, we also consider a case with a constant covariance matrix $\Sigma_t = \Sigma_0$, which we label B_{nd}. Both the SV-B_{nd} and B_{nd} specifications can be estimated quickly and robustly. The case with constant volatility is closely related, and hence can be compared with the widely used NA implementation of the model in Black (1995) by Wu and Xia (2016).

2.4 Estimation Strategy

A central practical objective of our paper is to introduce a set of models that are both accurate but also quick to estimate. An important insight in Joslin, Singleton, and Zhu (2011) is that implementation of affine models is easier if the candidate model is rotated to an observationally equivalent representation in which the state variables are the first principal components of yield or forward rates such that the state variables become observable.

The same insight is applicable to shadow rate models even if they do not belong to the affine class of models studied by Joslin, Singleton, and Zhu (2011). In the online appendix, we show that we can essentially undo the non-linearity of our model and transform it back into a linear space for the purpose of estimation. In the case of shadow rate models, the transformation recovers a shadow short rate and shadow forward rates that we use as observable risk factors with linear dynamics. Then, based on this representation, we provide the closed-form conditional likelihood of the data. Finally, we analytically concentrate out parameters of the VAR(1) from the likelihood, which preserves another appealing feature of the approach in Joslin, Singleton, and Zhu (2011). The fact that our construction for Σ_t uses $\Delta X_t \Delta X_t'$ to update the volatility is an important building block for this approach. In this way, the conditional mean (K_0, K_1) parameters in Equation (8) do not enter the volatility dynamics in Equation (9) and thus can be optimally obtained as a GLS estimate of the VAR(1).

Overall, we can achieve fast, convenient and robust estimation of every model. For

robustness, we also check that our key results are unchanged if we assume that the pricing portfolios are measured with error and use a Kalman filter to derive the likelihood. Note that a quick and robust estimation procedure is available more generally as long as the short rate function $m(.)$ is invertible and the $g(.)$ function remains affine. For even more general specifications of these functions, one could use the robust estimation method proposed by Andreasen and Christensen (2015) for non-linear dynamic term structure models with potentially non-Gaussian latent factors.

3 Model Evaluation in a Simulated Environment

We first estimate and evaluate several bond pricing models using a large number of simulated yield data sets. Using simulated yields makes comparisons between models more accurate because all true predictive moments, such as yield forecasts, volatility forecasts, and conditional Sharpe ratios, are known at each point in time in every simulation. This is unlike using historical data (which we do in the next section), where predictive moments are not known and some are difficult to estimate precisely. This is especially true if we want to carry the assessments to periods over which yields are close to the lower bound.

The data generating process for the simulations is the linear-rational term structure model of Filipovic, Larsson, and Trolle (2017) that satisfies the lower bound and allows for time-varying volatility. We select simulations with 30 years of data and that include an episode at the lower bound toward the end of the sample. The online appendix provides more details and shows that this sophisticated environment produces realistic simulated paths for yields and captures the upward sloping yield curve as well as the downward sloping yield volatility curve. This environment also produces realistic conditional Sharpe ratios that we can use to assess term structure models. Table 1 reports quantiles of the realized Sharpe ratios in U.S. data for investment horizons of 3, 6, and 12 months for bonds with 2, 5, and 10 years to maturity. For each bond and each investment horizon, we compute the realized Sharpe ratios by dividing the average excess return by its standard deviation over rolling windows of 36 months. Our rolling calculations attempt to capture, in a model-free way, the time variation in Sharpe ratios. Importantly, the same calculations can be applied to both historical yields and simulated yields to facilitate comparison.

Overall, while the estimates in Filipovic et al. (2017) were not targeted to match moments of the historical Sharpe ratios, the simulation closely matches the realized moments. Because yields have declined and average returns were higher after 2008, the historical median estimates

		Simulation					U.S. Data				
m	h	2.5	25	Median	75	97.5	2.5	25	Median	75	97.5
	2y	-0.51	-0.14	0.03	0.20	0.54	-0.64	-0.08	0.35	0.68	1.64
3m	5y	-0.50	-0.12	0.07	0.24	0.59	-0.49	-0.06	0.17	0.41	0.80
	10y	-0.49	-0.10	0.09	0.26	0.62	-0.50	-0.05	0.11	0.30	0.61
	2y	-0.76	-0.20	0.06	0.31	0.84	-0.97	-0.17	0.39	0.88	2.27
6m	5y	-0.76	-0.17	0.12	0.38	0.93	-0.81	-0.09	0.24	0.56	1.16
	10y	-0.74	-0.15	0.14	0.41	0.96	-0.88	-0.09	0.15	0.43	0.87
	2yr	-1.13	-0.28	0.12	0.52	1.36	-1.38	-0.22	0.38	1.16	2.49
12m	5y	-1.20	-0.24	0.21	0.63	1.51	-1.57	-0.12	0.32	0.75	2.07
	10y	-1.18	-0.22	0.25	0.67	1.61	-1.66	-0.17	0.24	0.68	2.00

Table 1: Sharpe Ratios Using Historical and Simulated Data

For each bond maturity m and each investment horizon h, we compute the realized Sharpe ratios by dividing the average excess return by its standard deviation over rolling windows of 36 months. The same calculations are applied to both historical yields and simulated yields. Columns (3)-(7): quantiles of the Sharpe ratios across all simulated samples. Columns (8)-(12): quantiles of the Sharpe ratios for the U.S. data.

are higher than the comparable statistics reported by Duffee (2010). The ex-post realized Sharpe ratios are higher than the conditional Sharpe ratios that actually prevailed in bond markets; this is likely because a substantial share of that decline in yields over the last few decades was unexpected. This may explain why the historical estimates tend to be somewhat higher than the estimates from simulations.

With simulations in place, we assess the performance of a number of models. Specifically, we consider an NA affine Gaussian model (G_{na}), an NA affine stochastic volatility model ($A1_{na}$), an NA Black shadow rate model (B_{na}) as in Wu and Xia (2016), a tractable ND Black shadow rate model (B_{nd}), and a tractable ND Black shadow rate model with stochastic volatility (SV-B_{nd}). For comparability, all models feature three pricing factors. The G_{na} and the $A1_{na}$ models correspond respectively to the well-known $A_0(3)$ and $A_1(3)$ models of Dai and Singleton (2000). The difference between each of the considered models and our adopted data generating process reflects the inescapable reality that the modeler will never know with certainty the true data dynamics. As an aside, while we focus on a linear-rational model to assess these models, we note that we find similar results using the Gaussian quadratic term structure model as the data generating process.

3.1 Sharpe Ratios

We use estimates of the conditional Sharpe ratios to evaluate and rank models. For a given investment horizon, the ratio's numerator is the conditional expected return in excess of the risk-free rate. The denominator is the conditional volatility of the corresponding return. Intuitively, the Sharpe ratio measures the expected return per unit of risk—where risk is measured by total volatility. Our ranking presumes that term structure models that accurately estimate conditional Sharpe ratios are more useful than other models that do not. For completeness, the online appendix reports results based on the estimates of the conditional mean and volatility of yields across models.

To establish this ranking, we consider three investment horizons up to 1-year ahead and four bond maturities up to 10 years. We measure the accuracy with the root median squared errors (RMedSE) between true and model-implied Sharpe ratios. We use the median both along the time dimension and across simulations. Using median statistics mitigates the risk that one extreme observation might skew the performance of any of the models.

Table 2 reports the RMedSE statistics implied by all models considered. The columns labeled μ and σ report the mean and standard deviations of Sharpe ratio estimates based on estimates of bond returns and volatility computed using rolling windows of 36 months. The column labeled T reports the RMedSE based on the rolling-window Sharpe ratio estimate, thereby providing a practical and realistic baseline against which to gauge the other models. The left side of Table 2 reports statistics computed using the full sample in every simulation. The right side of the table reports statistics using only sub-samples close to the lower bound.

Looking at the results, we ask two questions. First, how do the NA and ND implementations of the Black shadow rate model compare with each other? The answer is that the B_{nd} model delivers more accurate Sharpe ratio estimates than the B_{na} model, on average, for every investment horizon and bond maturity that we consider, without exception. In fact, the B_{na} model sometimes even underperforms the simpler Gaussian G_{na} model. This contrast is rather stark over the lower-bound sample. Whereas the B_{nd} model handily outperforms the model G_{na}, the B_{na} model performs consistently worse than the G_{na} model, barring one exception. We discuss the causes underlying these differences in the next sub-section.

The second question we ask is what is the relative gain from modeling both the lower bound and time-varying volatility of yields? The answer to this question is that the ND Black model with stochastic volatility, SV-B_{nd}, significantly outperforms other models. In full samples, the SV-B_{nd} model delivers the most accurate Sharpe ratios, on average, for all

		Entire Sample								ZLB Sample							
h	m	μ	σ	T	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	μ	σ	T	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
3m	6m	0.02	0.26	0.27	0.22	0.22	0.36	0.27	0.18*	0.18	0.18	0.51	0.29	0.30	0.28	0.17*	0.21
	2y	0.04	0.26	0.38	0.47	0.46	0.76	0.48	0.36*	0.22	0.17	1.00	0.64	0.69	0.68	0.33	0.33*
	5y	0.07	0.27	0.57	0.83	0.78	1.15	0.75	0.58*	0.23	0.19	1.28	1.03	0.99	1.16	0.70	0.47*
	9y	0.08	0.27	0.73	1.17	1.09	1.40	1.07	0.84*	0.23	0.21	1.41	1.26	1.32	1.27	1.03	0.77*
6m	6m	0.05	0.38	0.34	0.18	0.18	0.32	0.24	0.16*	0.29	0.27	0.53	0.20	0.20	0.23	0.14*	0.17
	2y	0.07	0.40	0.45	0.41	0.40	0.68	0.39	0.32*	0.35	0.27	0.92	0.48	0.49	0.61	0.27	0.24*
	5y	0.11	0.42	0.65	0.73	0.70	1.05	0.67	0.50*	0.38	0.29	1.23	0.75	0.83	0.97	0.58	0.38*
	9y	0.13	0.42	0.78	1.07	0.99	1.26	0.97	0.75*	0.37	0.31	1.38	1.06	1.12	1.16	0.97	0.70*
12m	6m	0.10	0.59	0.46	0.16	0.14	0.25	0.19	0.14*	0.52	0.43	0.69	0.16	0.14	0.18	0.12*	0.13
	2y	0.15	0.64	0.60	0.35	0.33	0.55	0.34	0.27*	0.62	0.45	1.00	0.35	0.32	0.47	0.25	0.19*
	5y	0.21	0.67	0.80	0.64	0.60	0.93	0.61	0.42*	0.65	0.46	1.26	0.65	0.64	0.82	0.48	0.33*
	9y	0.23	0.68	0.88	0.93	0.85	1.15	0.90	0.63*	0.65	0.49	1.44	0.91	0.85	1.02	0.86	0.55*

Table 2: Estimating Conditional Sharpe Ratios in Simulations

Precision of the conditional Sharpe ratio estimates at horizon h (column 1) and maturity m (column 2). μ and σ refer to the median across simulations of the mean and standard deviations of rolling-window Sharpe ratio estimates. The other columns report the root median squared errors (RMedSE) between true and estimated conditional Sharpe ratios. T refers to RMedSE of rolling-window Sharpe ratio estimates. G_{na}, $A1_{na}$, and B_{na} refer to the NA affine Gaussian model, NA affine stochastic volatility model, and the NA shadow rate model, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively. The symbol * indicates the best performance for a given horizon, maturity, and sample. The lower-bound sub-samples collect dates when the one-month rate is 25 basis points or less.

considered horizons and bond maturities. The RMedSE gains are sizeable relative to the baseline rolling-window estimates and relative to the NA Gaussian model, ranging from 10 to 30 percent. Finally, the SV-B_{nd} gains are smaller relative to the B_{nd} model, but larger relative to the B_{na}. The $A1_{na}$ model also offers gains in full samples.

When focusing on the lower-bound samples, the baseline rolling-window estimates offer much poorer accuracy relative to the full-sample results. However, the SV-B_{nd} model still delivers the best performance for all horizon-maturity combinations, with the exception of one case where the B_{nd} is ranked first (6-month horizon and 6-month terminal maturity). In all cases, the RMedSE gains are large relative to either of the NA models, ranging between 15 to 50 percent.

The winning model in Table 2 is always statistically significantly better than the G_{na} model. To see this clearly, we can look at the distribution of simulation results. Figure 2 reports the cumulative empirical distribution of RMedSEs across simulations. For clarity, we include results for the SV-B_{nd}, G_{na}, and B_{na} models (the distribution for the $A1_{na}$ model is close to the G_{na}, like their respective medians in Table 2). This distribution accounts for the sampling variability due to the sample size, to the model estimation, and to mis-specification.

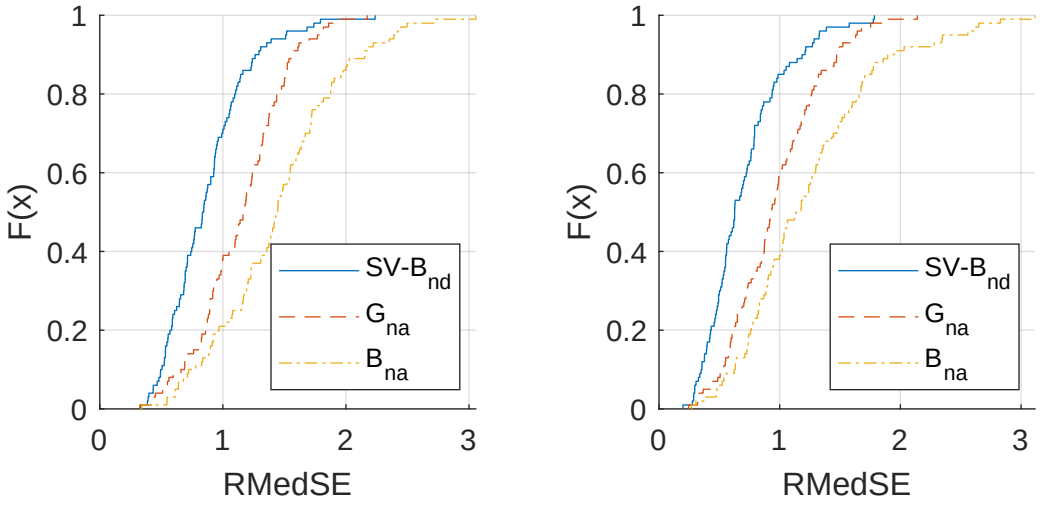
The ranking between the three models clearly emerges. The SV-B_{nd} exhibits first-order stochastic dominance over the benchmark G_{na}, which itself exhibits first-order stochastic dominance over the B_{na} model. This means that any quantile for RMedSe in the distribution from the SV-B_{nd} model is lower than in the distribution from the G_{na}. A test for equality also rejects the null hypothesis for arbitrary small critical values. As mentioned earlier, we find similar results (unreported) when using the Gaussian quadratic term structure model as the data generating process.

We also evaluate the accuracy of Sharpe ratios for *portfolios* of bonds, which depends on the models' ability to forecast the correlations among future yields. We form three simple equal-weighted portfolios consisting of zero coupon bonds with maturities of 6 months and 1, 2, ..., 9 years. The first portfolio combines all 10 maturities, the second portfolio combines the first 5 short-term maturities and the third portfolio combines the last 5 maturities. Table 3, sharing the same column structure as Table 2, reports the performance by the four models in forecasting the Sharpe ratios of the three portfolios. The key messages are essentially unchanged. The SV-B_{nd} model clearly outperforms all other models for every category. The performance gains obtained by the SV-B_{nd} model, relative to the Gaussian model, are similar to those observed for the case of the individual bonds reported in Table 2. In particular, and as expected, the gains are especially large across samples near the lower bound.

		Entire Sample					ZLB Sample				
h	m	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
	6m-9y	0.5	0.97	1.51	0.94	0.72*	0.7	0.99	1.04	0.49	0.45*
3m	6m-4y	0.9	0.92	1.39	0.95	0.68*	1.0	0.99	1.10	0.73	0.49*
	5y-9y	1.0	0.91	1.30	0.92	0.68*	1.2	1.04	1.04	0.76	0.52*
	6m-9y	0.5	0.96	1.60	0.98	0.77*	0.6	0.99	1.20	0.58	0.46*
6m	6m-4y	0.8	0.95	1.43	0.94	0.69*	0.8	1.06	1.22	0.81	0.53*
	5y-9y	0.9	0.95	1.30	0.94	0.68*	0.9	1.06	1.15	0.86	0.58*
	6m-9y	0.4	0.96	1.61	1.03	0.77*	0.4	0.91	1.39	0.72	0.52*
12m	6m-4y	0.7	0.94	1.45	1.00	0.66*	0.7	0.96	1.24	0.79	0.53*
	5y-9y	0.8	0.94	1.31	1.01	0.66*	0.8	0.99	1.27	0.95	0.58*

Table 3: Bond Portfolio Sharpe Ratios

G_{na}, $A1_{na}$, and B_{na} refer to the NA affine Gaussian model, the NA affine stochastic volatility model, and the NA shadow rate model, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively. For the G_{na} model, we report the root median squared errors (RMedSE) between true and model-implied estimates. For other models, we report the RMedSE statistics relative to the G_{na} model. The first two columns give the forecast horizon h and the maturity m at the end of the investment periods, respectively. The symbol * indicates the best performance for a given forecast horizon, bond maturity, and sample. The lower-bound sub-samples collect dates for which the one-month rate is 25 basis points or less.



(a) $m = 9$ years and $h = 3$ months (b) $m = 9$ years and $h = 12$ months

Figure 2: RMedSE Distributions

Cumulative distribution function of simulated RMedSEs from Sharpe ratio estimates. G_{na} and B_{na} refer to the NA affine Gaussian and Black models, respectively. SV-B_{nd} refers to the ND Black models with BEKK volatility. Both panels report results for holding a bond with nine years to maturity at the end of the investment horizon. Panel (a): results for the 3-month holding horizon. Panel (b): results for the 12-month holding horizon.

Discussion of the Simulation Results

The combination of time-varying volatility and satisfying the lower bound appears important in driving the performance in these simulations. In the online appendix, we show that the flexible volatility dynamics help improve forecasts of the denominator in the Sharpe ratio (which is expected) as well as the numerator (which is more surprising). The latter gains are due to the increased econometric efficiency.

It is easy to understand why the SV-B_{nd} model does a better job, given its flexible volatility dynamics. But what can explain the poor performance of the NA Black model B_{na} relative to the ND Black model B_{nd}? Fortunately, the simulation exercise can provide further insights into the differences.

Conceptually, the only difference between these models can be found in the pricing equations, which is where the answer must lie. To help facilitate comparison, the pricing

equation in the B_{na} model is given by:

$$f_{n,t} \approx S_n \, w \left(\frac{A_n + B_n' X_t}{S_n} \right),$$

where A_n and B_n are restricted by the absence of arbitrage (the coefficients S_n, A_n, and B_n are given in the online appendix). Some of the parameters that constitute S_n play a dual role in the B_{na} models. They govern bond pricing and contribute to fitting the volatility of yields. This dual role may create tension between fitting yields and volatilities. Notably, this tension is largely absent in ND models (see Equation 7).[6]

Any trade-off between yield and volatility forecasts will affect the Sharpe ratio estimates. However, poorer yield and volatility estimates do not always lead to worse Sharpe ratio estimates. To illustrate this possibility, consider an investment with an expected excess return and expected volatility of 0.3 and 0.1, respectively, where the true Sharpe ratio is 3. An error of +0.1 in the forecasts of returns and volatilities implies a conditional Sharpe ratio of $0.4/0.2 = 2$. This is a difference of 1 relative to the true ratio. By contrast, an error of -.05 in the forecasts of the expected returns and return volatility implies a conditional Sharpe ratio of $0.35/0.05 = 5$. This is now a difference of 2 relative to the true ratio. The smaller absolute forecast errors produce a larger Sharpe ratio error.

Following this illustration, and to be more precise, we can express the mean squared error of Sharpe ratios analytically. Define the forecast error $fe_j(x)$ and the mean squared forecast error $msfe_j(x)$ for some forecast x from model j:

$$\begin{aligned} fe_j(x) &\equiv (x_j - x_0) \\ msfe_j(x) &\equiv E[fe_j(x)^2 | x_0] = E[(x_j - x_0)^2 | x_0], \end{aligned}$$

where we suppress the horizon and maturity index for clarity, x_j is the model forecast from model j, and x_0 is the true forecast in the simulation. For instance, $msfe_j(xr_{t,t+h}^{(m)})$ gives the mean squared return forecast error for the yield with maturity m at the horizon h, and $msfe_j(\sigma_{t,t+h}^{(m)})$ gives the mean squared volatility forecast error. Analytically, it can be shown

[6]The scaling parameter θ of the B_{nd} model does not play a role in shaping the volatility dynamics of states. However, we require that $\theta \, w\left(\frac{0-\sigma_s}{\theta}\right) \geq 0.0005$ and $\theta \, w\left(\frac{0+\sigma_s}{\theta}\right) \leq 0.0050$. These inequalities constrain the variation of the shadow rate only locally, near the lower bound. The NA restrictions tie S_n at every maturity with the underlying volatility parameters.

that $msfe_j(SR^{(m)}_{t,t+h})$ is given by:

$$\begin{aligned} msfe_j(SR) \approx\ & msfe_j(xr) + b_1 msfe_j(\sigma) \\ & + b_2 E\big[fe_j(xr)\big] \times E[fe(\sigma)\big] \\ & - b_3 COV\big(fe_j(xr); fe_j(\sigma)\big), \end{aligned} \tag{10}$$

where the coefficients are positive $b_1 > 0$, $b_2 > 0$ and $b_3 > 0$ and they depend on the true Sharpe ratio (see the online appendix). This expression is useful to confirm the intuition above. The first term indicates that the accuracy of Sharpe ratio estimates moves one-for-one with the accuracy of returns forecasts. The second term indicates that the impact from the accuracy of the volatility forecasts depends on the level of the Sharpe ratio. The third term indicates that $msfe_j(SR)$ also increases if the product of the biases is positive, which is caused by the non-linearity in the ratio. Finally, the last term shows that the $msfe_j(SR)$ increases with the correlation between the return forecast error and the volatility forecast error $fe_j(xr)$ and $fe_j(\sigma)$.

In the online appendix, we show that the forecast biases are the key drivers for the relatively poor performance of the NA models in the simulation exercise. Both the B_{na} and G_{na} models exhibit large biases in their yield and volatility forecasts. In the case of the B_{na} model, both biases are negative and reduce the accuracy of the Sharpe ratio estimates. For the G_{na} model, the biases have opposite signs, which mitigates the effect on the accuracy of the Sharpe ratio estimates. We also show that the RMSEs of yield and volatility forecasts do not explain the relatively poor performance of NA models in estimating conditional Sharpe ratios. Unsurprisingly, the SV-B_{nd} model delivers the most accurate forecasts with the smallest biases across the simulation results.

4 Empirical Illustrations

In this section, we complement our simulation evidence and examine the performance of the G_{na}, $A1_{na}$, B_{na}, B_{nd} and SV-B_{nd} models based on historical data.[7] We use rates,

[7]The linear-rational model of Filipovic, Larsson, and Trolle (2017), used as the data generating process in Section 3, is capable of capturing many desirable features of the data. However, its estimation is significantly more involved and requires auxiliary information from derivatives markets, which they use in their paper. Hence, given our focus on tractability, we do not report empirical results based on the linear-rational model. To maintain a level playing field, we include only those models for which standard estimation is based on yield data only and whose global convergence can be obtained in a straightforward and robust manner.

measured monthly from January 1970 through August 2020, on 1-month forward loans with 13 different maturities: 1-month, 3-month, 6-month, 1-year, 2-year, and annually until year 10. Specifically, we use the data provided by Gurkanyak, Sack, and Wright (2007) for maturities longer than 6 months. For maturities less than or equal to 6 months, we bootstrap the forward rates from Treasury bond prices provided by the Center for Research in Security Prices (CRSP).

Based on these historical data, we report pricing errors, yield and volatility forecasts and shadow rate estimates, relative to the G_{na} model. Following Wu and Xia (2016), and for comparability with existing results, we use the Kalman filter estimate of the B_{na}, B_{nd} and the G_{na} models in our analysis. Consistent with the evidence in Joslin, Le, and Singleton (2013), our results are unchanged if we estimate every model with shadow forward portfolios priced without error. For the $A1_{na}$ and SV-B_{nd} models, we use the estimate obtained by assuming the portfolios of shadow forwards are priced perfectly. This approach is easy to implement and converges quickly and robustly to global parameter estimates.

4.1 Pricing Errors

Table 4 reports the RMSEs in fitting yields for selected maturities in the full sample and in the sub-sample where the short rate is at the lower bound. First, the results are very similar across models. This is not surprising, because every model that we consider has three factors, which is well known to be enough to fit the cross-section of yields.

	Entire Sample					ZLB Sample				
m	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
6m	22.2	21.1	18.2	17.7	21.3	15.1	14.8	13.7	12.1	16.1
1y	18.7	18.5	15.2	15.3	18.4	11.9	11.1	7.2	7.6	10.9
2y	12.4	12.7	10.4	10.6	12.1	9.8	10.3	4.5	5.1	6.7
5y	5.2	3.2	4.0	3.9	3.5	4.5	2.0	3.0	2.7	3.3
10y	1.9	1.3	1.4	1.4	1.4	1.4	0.9	0.9	0.7	1.1

Table 4: Pricing Errors

Pricing errors RMSEs for NA and ND models (in basis points). G_{na}, $A1_{na}$, and B_{na} refer to the NA affine Gaussian model, the NA affine stochastic volatility model, and the NA shadow rate model, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively. Sample period is 1970:Jan - 2020:Aug.

Second, we contrast the results between the NA and ND shadow rate models with constant

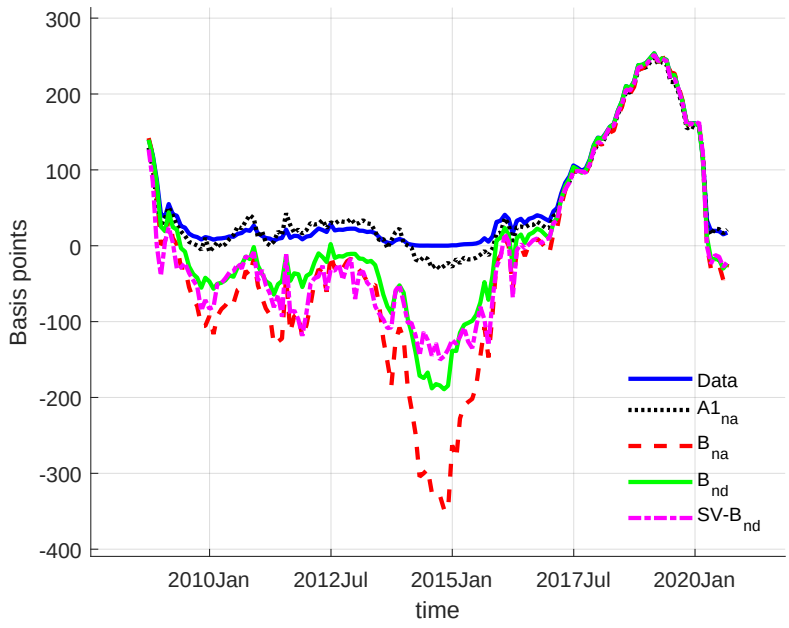


Figure 3: Shadow Rates

One-month short rate and shadow rate (1970:Jan–2020:Aug). $A1_{na}$ and B_{na} refer to the NA affine stochastic volatility and the NA shadow rate models, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively.

volatility, B_{na} and B_{nd}. We find that the pricing errors are very close, within a fraction of 1 basis point (bp) over the full sample. This result parallels existing work finding negligible differences between the pricing errors of the Gaussian G_{na} models and those of the DNS models. Table 4 extends this result to models with a Black truncated short rate. Note that both models offer a small improvement over the standard Gaussian G_{na} model. As expected, this difference is concentrated in the period close to the lower bound.

4.2 Shadow Short Rates

Figure 3 plots the one-month rate in the $A1_{na}$ and the one-month shadow rates implied by the Black models B_{na}, B_{nd} and SV-B_{nd}, along with the observed one-month short rate, focusing on the period after 2008. For much of the first lower-bound period, it is rather interesting that the shadow rates implied by the two ND models, despite their different volatility specifications, remain close to one another. Overall, the NA model B_{na} shadow rates remain within a range of 50 bps with one clear exception: for much of 2014, the shadow rate implied by the NA model B_{na} falls much deeper into the negative region, often by more

than 100 bps, relative to the two ND shadow rates. In the more recent lower-bound period, starting in March 2020, these three models produce similar shadow rate estimates. The estimated shadow rate for the B_{na} model that we report is broadly similar to results reported by the Federal Reserve Bank of Atlanta based on the Wu and Xia (2016) model. In particular, both also report a shadow rate at -3 percent in 2014.

The pattern in Figure 3 suggests that while there has been considerable interest in studying the economic content of shadow rates, this must be done with caution given that shadow rates can be highly model-dependent. This is also an important observation documented in Bauer and Rudebusch (2016).

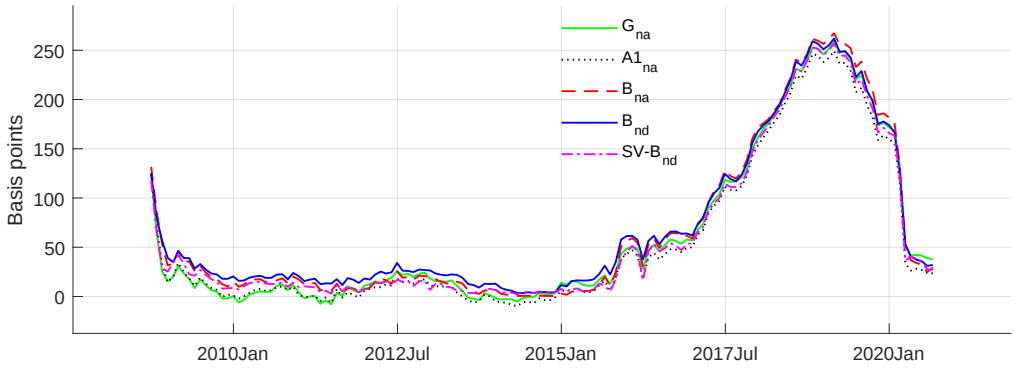
4.3 Short Rate Forecasts

Figure 4 shows each model's forecast of the 1-month rate, focusing on the period close to the lower bound after 2008. We report results for the 3- and 12-month forecast horizons. Results for other maturities lead to similar observations (not reported).

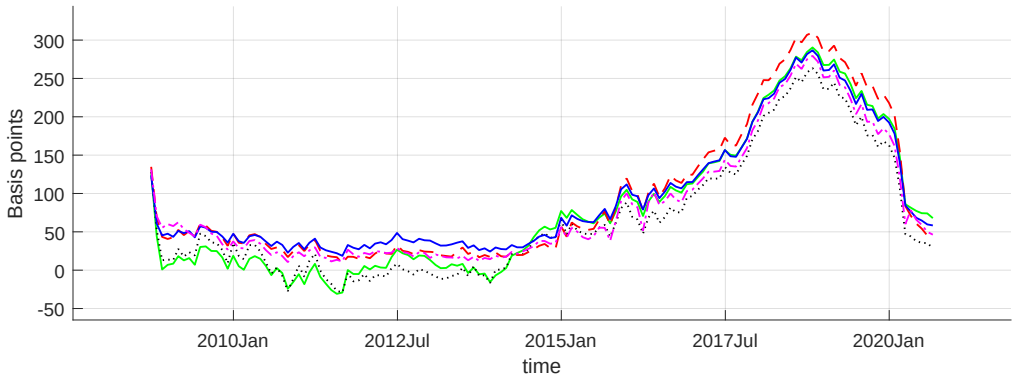
Two observations are worth nothing. First, the forecasts implied by the NA affine Gaussian model (green lines) and the NA stochastic volatility model (dotted line) can be visibly distinct from the forecasts implied by the shadow rate models. In particular, their forecasts significantly violate the lower bound (at all horizons) in the prolonged period near the lower bound after 2008. Second, regardless of the model type—NA, ND, constant variance or stochastic volatility—the forecasts by the shadow rate models are close to one another. Their proximity suggests that it can be challenging to accurately compare the different models in one short sub-sample close to the lower bound. This observation strengthens the case for the exercise in Section 3, where the models are evaluated in a controlled simulation environment. Additionally, this observation serves as a reminder that even large differences in the model-implied shadow rates, as documented in the previous sub-section, should be interpreted with caution. As we can see here, large differences in shadow rates do not translate to empirically meaningful differences in short rate forecasts.

4.4 Volatility Forecasts

It might be challenging to compare models based on the forecast of yields in one short sample close to the lower bound, as we have argued above. But this argument does not carry fully to comparison based on the models' implied yield volatility. Figure 5 plots the volatility forecasts of the 24-month yield implied by the models over two forecast horizons: 3- and



(a) 3-month forecast horizon



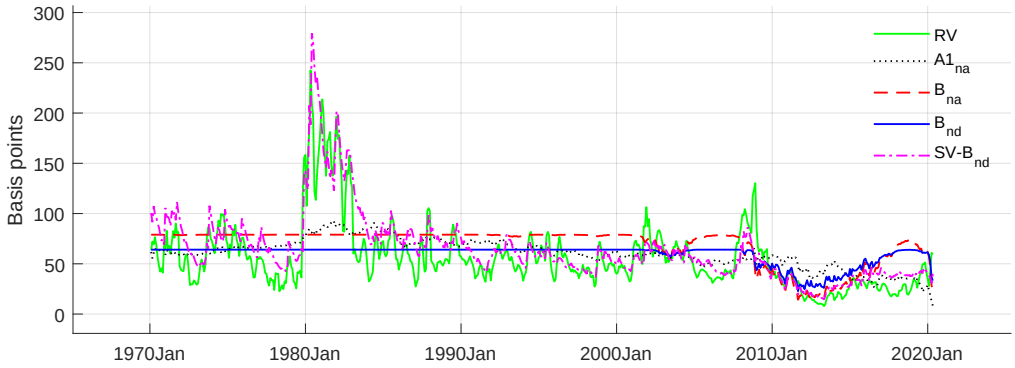
(b) 12-month forecast horizon

Figure 4: Short Rate Forecasts

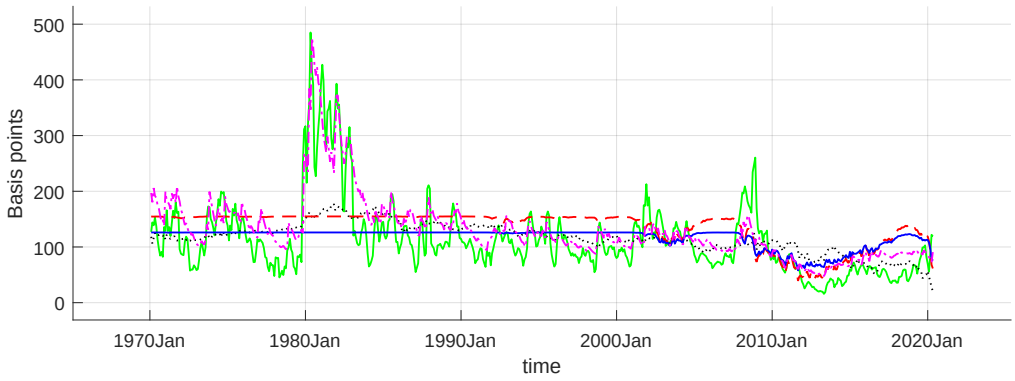
Forecasts of the one-month rate (1970:Jan–2020:Aug). G_{na}, $A1_{na}$, and B_{na} refer to the NA affine Gaussian model, affine stochastic volatility model, and shadow rate model, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively.

12-month. As a benchmark, we compute the realized variances of the 24-month yield using three months of daily data leading up to each point in time t. It is known that realized volatility can be estimated precisely using relatively high frequency data, even in a short window.

Every shadow rate model appears to do a good job during the lower-bound episodes starting in 2008 and 2020. Away from the lower bound, the SV-B_{nd} model performs well matching the large fluctuations of the RV series (ranging from as low as 40 bps to as high as 300 bps). This justifies the simple construction for the volatility dynamics in Equation (9). The volatility forecasts from the $A1_{na}$ model follow the broad pattern displayed by the



(a) 3-month forecast horizon



(b) 12-month forecast horizon

Figure 5: Volatility Forecasts

Forecasts of the 2-year bond yield volatility (1970:Jan–2020:Aug). G_{na}, $A1_{na}$, and B_{na} refer to the NA affine Gaussian model, affine stochastic volatility model, and shadow model, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively.

realized volatility. Specifically, the $A1_{na}$ model tracks rather closely the realized volatility pattern in the recent episode near the lower bound and continues to do so when yields are lifted off the lower bound. However, the $A1_{na}$ volatility forecasts significantly miss the short period of very high volatility in the early 1980s. These results are consistent with the conclusion by Jacobs and Karoui (2009) that the ability of this class of models to capture conditional volatility depends on the sample period.

In contrast, and as expected, it is clear that the yield volatilities implied by the constant-variance Black models are essentially a flat line for periods away from the lower bound. Loosely speaking, the flat lines correspond to the average volatility levels away from the lower

bound implied by the B_{na} and B_{nd} models. It is notable that they are not at the same level. Specifically, the average volatility level implied by the NA model B_{na} is higher than the level implied by the ND model B_{nd} and higher than the corresponding RV series for much of the sample, with the exception of the early 1980s. The fact that the B_{na} model seems to miss the average volatility level indicated by the RV series constitutes evidence that the B_{na} model might be constrained. As we argue before, the scaling parameter S_n in the pricing equation of the B_{na} model is directly linked to the volatility parameters of the model. As such, this dual role of the scaling parameter might create a tension in the ability of the model to fit the multiple dimensions of the data.

These results are confirmed by the RMedSEs reported in Table 5. Across maturities and horizons, the SV-B_{nd} model is able to reduce volatility forecast errors by 10 to 30 percent relative to the $A1_{na}$ model during the full sample. The improvement can be even larger in specific cases or in sub-samples. The reduction is particularly large for short-term yields near the lower bound, where the improvement ranges from 70 to 85 percent. This is precisely where we expect the SV-B_{nd} model to provide the largest improvement. Across other maturities and in the period near the lower bound, the improvement ranges from 10 to 50 percent. Comparing the B_{na} and B_{nd} models during the period near the lower bound, where these models exhibit meaningful volatility dynamics, does not produce a clear winner.

		Entire Sample				ZLB Sample			
h	m	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
	1m	35.85	1.15	0.87	0.76	34.49	0.61	0.50	0.16
	2y	16.52	1.48	1.04	0.69	15.28	0.73	1.13	0.50
3m	5y	12.00	1.43	1.18	0.83	8.77	1.23	1.40	0.79
	10y	12.34	1.25	1.09	0.71	13.65	0.96	0.80	0.47
	Avg	19.17	1.33	1.05	0.75	18.05	0.88	0.96	0.48
	1m	49.73	1.10	0.85	0.77	46.07	0.68	0.61	0.20
	2y	22.61	1.50	1.08	0.73	20.89	0.79	1.21	0.66
6m	5y	16.87	1.39	1.17	0.86	13.24	1.13	1.24	0.84
	10y	17.64	1.22	1.07	0.74	19.60	0.92	0.82	0.51
	Avg	26.71	1.30	1.04	0.77	24.95	0.88	0.97	0.55
	1m	67.85	1.05	0.86	0.71	60.66	0.88	0.73	0.31
	2y	30.34	1.52	1.15	0.89	28.00	0.99	1.35	0.92
12m	5y	23.05	1.39	1.17	0.92	19.78	1.04	1.18	0.88
	10y	25.35	1.17	1.05	0.76	28.33	0.90	0.76	0.57
	Avg	36.65	1.28	1.06	0.82	34.19	0.95	1.00	0.67

Table 5: Volatility Forecasts

Root median squared difference between volatility forecasts and our realized volatility forecasts for a bond with maturity m and an horizon h. $A1_{na}$ and B_{na} refer to the NA affine stochastic volatility model and the NA shadow rate model, respectively. B_{nd} and SV-B_{nd} refer to the ND models with and without BEKK volatility, respectively. For the B_{na} model, we report the root median squared errors (RMedSE), in basis points, between RV and model-implied volatility forecasts. For the B_{nd} and SV-B_{nd} models, we report the ratio of RMedSe relative to the $A1_{na}$ model. Forecasts horizons are 3, 6, and 12 months.

5 Model Evaluation with Out-of-Sample Forecasts

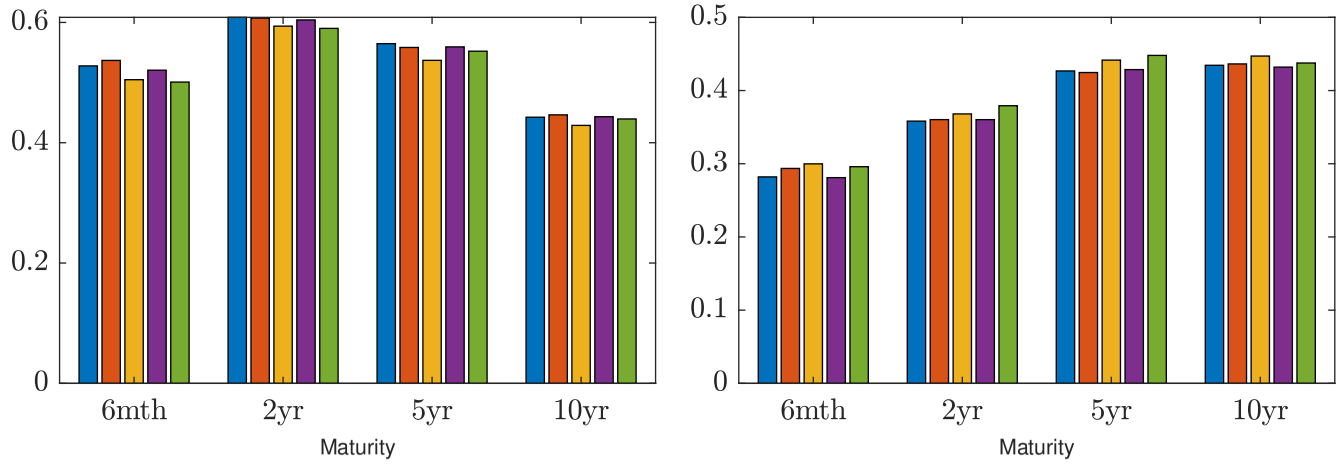
This section continues the evaluation of the G_{na}, $A1_{na}$, B_{na}, B_{nd}, and SV-B_{nd} models in a real-time, out-of-sample forecast exercise based on the same historical data employed in the previous section. To start, each model is first estimated based on monthly historical data from January 1970 through December 1995. As above, we use rates on 1-month forward loans with 13 different maturities: 1-month, 3-month, 6-month, 1-year, 2-year, and annually until year 10. We then evaluate 3-month-ahead yield and volatility forecasts for the 6-month, 2-year, 5-year, and 10-year bonds, and we record the associated forecast errors. The results are qualitatively similar if we look at 6-month or 12-month forecasts. We then re-estimate each model expanding the sample to January 1996 data, and repeat the evaluation. We iterate this procedure, expanding the sample one month at a time, until August 2020. Figure 6 reports out-of-sample root median squared forecast errors for yields and volatilities across all time periods.

5.1 Yield and Volatility Forecasts

We start with an evaluation of yield forecast errors. Panels (a)-(b) of Figure 6 report the results in sub-samples before 2008 and after 2008 (away or near the zero lower bound), respectively. Despite the heterogeneity that we observe in the simulation environment, we do not observe significant differences across the models in terms of their ability to characterize yield expectations out-of-sample. Regardless of whether we are away from or near the lower bound, out-of-sample yield forecasts with an expanding estimation window do not seem to be a metric for which model delineation is particularly useful.

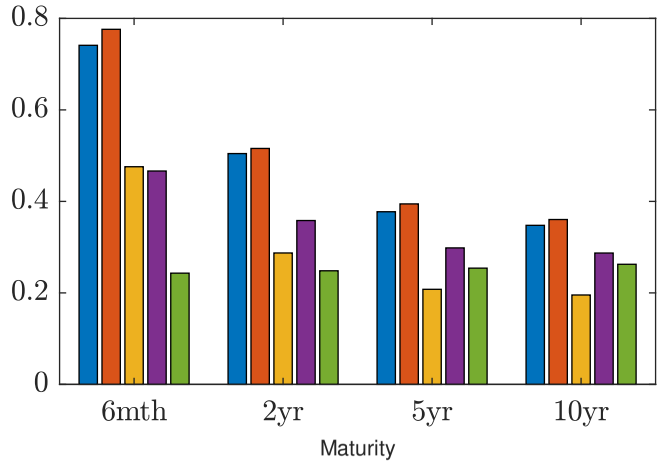
We next turn to an evaluation of volatility forecasts, where the realized volatility is the square root of the sum of daily squared returns. Panels (c)-(d) of Figure 6 also report the results in sub-samples before 2008 and after 2008, respectively. Regardless of the proximity to the zero lower bound, the volatility forecasting performance of the Gaussian model, G_{na}, is, not surprisingly, relatively poor. The two best models in terms of volatility forecasting are, also not surprisingly, the $A1_{na}$ and SV-B_{nd} models. The SV-B_{nd} model produces volatility forecast errors that are consistently small, on average, across maturities, and its performance relative to other models is particularly notable for shorter maturity bonds. In contrast, the performance of the $A1_{na}$ model is strong overall, but deteriorates for shorter maturity bonds.

When we compare the performance of the B_{na} and B_{nd} models, we find that the ND model performs better before 2008, particularly for shorter maturity bonds. This difference

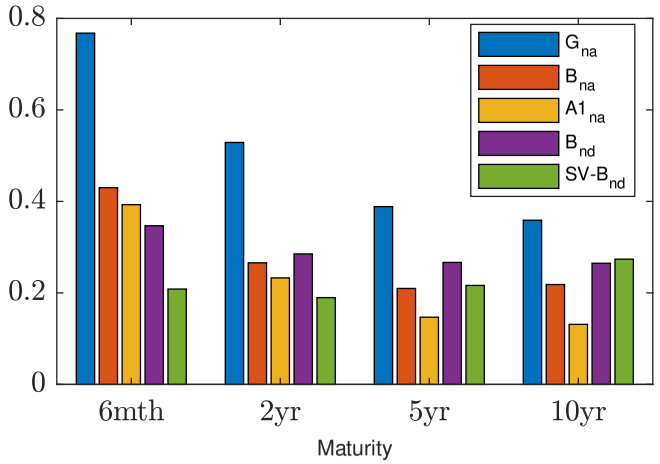


(a) Yield forecast RMSEs—Before 2008

(b) Yield forecast RMSEs—After 2008



(c) Volatility forecast RMSEs—Before 2008



(d) Volatility forecast RMSEs—After 2008

Figure 6: Out-of-sample Forecast Errors

Root mean squared errors (RMSEs) in 3-month-ahead out-of-sample forecasts of yields and yield volatilities. Forecasts are based on estimation from January 1970 through December 1995 and adding one observation every month until August 2020. G_{na}, $A1_{na}$, and B_{na} refer to the NA affine Gaussian model, the NA affine stochastic volatility model, and the NA shadow rate model, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively.

can only be attributed to the bias introduced by the no-arbitrage restriction. However, the performance of both improves after 2008 and the wedge between the two models essentially closes (Panel (d) of Figure 6). For long maturities, their performance is close to the results based on the SV-B_{nd} model. This latter result suggests that the volatility compression that arises near the lower bound is capturing most of the variation in volatility during these periods (see also Christensen and Rudebusch 2014; Kim and Priebsch 2013).

Overall, a real-time out-of-sample exercise based on yield forecasting proves less useful for delineating models. This is perhaps not so surprising given the extremely persistent, near unit-root nature of bond yields. This was also part of the motivation for the simulation exercise in Section 3. However, the exercise is able to delineate models based on volatility dimensions, highlighting the importance of the models that directly facilitate stochastic volatility. Taken together, our preferred SV-B_{nd} model fares favorably in an out-of-sample evaluation. From this perspective, the closest competitor (particularly for some longer maturities) is the $A1_{na}$ stochastic volatility model. Hence, the results in the historical sample offer a more positive message about this model than the results in simulated samples.

5.2 Linear Projection of Yield Changes

Beyond RMSE statistics, Dai and Singleton (2002) propose a modified version of the popular Campbell and Shiller (1991) expectations hypothesis regression as an alternative metric to assess the predictive power of a term structure model. The original Campbell and Shiller (1991) regression is often referred to as LPY(i) and the modified version LPY(ii). In using essentially de-trended variables such as yield changes and the slope of the yield curve, this approach affords an evaluation of each model's ability to predict yields while largely abstracting from the near unit root behavior of bond data.

As is well-known, LPY(i) involves regressing a series of yield changes on the corresponding slopes of the yield curve. The regression is designed such that if the expectation hypothesis holds, the coefficient of the regression should be one, regardless of yield maturity. What Campbell and Shiller (1991) find, however, is that these coefficients are significantly different from one. As a matter of fact, they are mostly negative and increasingly so as bond maturity increases. This pronounced failure of the expectations hypothesis translates to evidence for strong time variation of bond risk premia. LPY(ii) adds a model-implied risk premium component to the yield changes used in LPY(i). The premise is that if the risk premium adjustment is in agreement with the data, the regression coefficients will be corrected and moved closer to one.

We revisit the exercise to evaluate the G_{na}, $A1_{na}$, B_{na}, B_{nd}, and SV-B_{nd} models. As in Dai and Singleton (2002), we ask whether the LPY(ii)-adjustment induced by each model can move the Campbell and Shiller (1991) regression coefficients closer to unity. It should be noted that in the out-of-sample spirit, we conduct this analysis on a real-time basis, whereas Dai and Singleton (2002) employ model estimates based on the entire sample.

Specifically, LPY(ii) is derived from the decomposition of the bond return into a risk premium component and an expectations component. Based on this decomposition, one can use the model-implied premium component as a correction term for LPY(i):

$$y_{n-1,t+1} - y_{n,t} + \frac{e_t^n}{n-1} = \alpha_n + \phi_n(\frac{y_{n,t} - y_{1,t}}{n-1}) + \epsilon_{n,t+1}, \tag{11}$$

where e_t^n is the relevant risk adjustment:

$$e_t^n = -(n-1)\hat{E}_t[y_{n-1,t+1} - y_{n,t}] + \hat{y}_{n,t} - \hat{y}_{1,t}. \tag{12}$$

The correction term e_t^n can be computed for every model. Equation (12) tells us that the ability of a candidate model to capture the risk adjustment in the context of LPY(ii) depends upon its ability to fit bond yields contemporaneously and to capture the time-series yield dynamics. If that is the case, then the theoretical value of the ϕ_n coefficient is unity. Note that models with similar forecast RMSEs may offer different performance based on the LPY(ii) criteria—the estimates for ϕ_n can differ—especially if the yield forecast errors are correlated with the spread $y_{n,t} - y_{1,t}$.

To operationalize LPY(ii), we require yield forecasts as well as fitted yields from every model. The typical way to compute e_t^n is to estimate each model using the entire sample of data. However, given our focus on out-of-sample performance, we estimate each model every month in the expanding window framework discussed above. We then compute e_t^n every month, collect the results from January 1995 until the end of the sample and run the LPY(ii) in Equation (11). This means that the adjustment term is a real-time estimate without look-ahead bias.

Figure 7 plots the LPY(ii) coefficients ϕ_n across maturities n for every model. We consider an out-of-sample forecast horizon of 12 months. There are two key takeaways. First, similar to Dai and Singleton (2002), we find that the coefficients implied by the $A1_{na}$ model remain far from unity (confidence bounds are provided in the blue shaded area). The premium adjustments have a relatively small effect and these coefficients are much closer to the standard

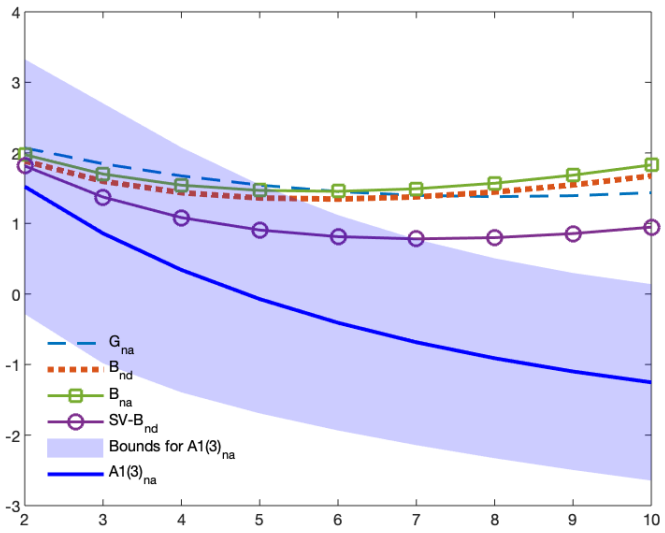


Figure 7: LPY(ii) Regression Coefficients

Estimates of coefficients ϕ_n in the LPY(ii) regression for annual yield changes based on Equations (11)-(12), where the adjustment e_t^n is computed based on real-time model estimation with an expanding sample. G_{na}, $A1_{na}$, and B_{na} refer to the NA affine Gaussian model, affine stochastic volatility model, and shadow rate model, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively. Confidence bounds for the LPY(ii) regression coefficients from the NA affine stochastic volatility model are provided in the blue shaded area.

Campbell-Shiller regression coefficients (i.e., ignoring the adjustment term). Despite the fact that the $A1_{na}$ performance in out-of-sample volatility forecasts is competitive, we infer that the well-known tension between yields, returns, and volatility dynamics in this no-arbitrage model has important negative consequences for the implications of out-of-sample forecasts in correctly capturing premium dynamics. Second, while there is relatively little separating the other models along this particular evaluation dimension, the coefficients implied by the SV-B_{nd} model are closer to unity across maturities than the $A1_{na}$ model. Taken together, from the perspective of yield and volatility forecasts along with LPY(ii) premium dynamics, the SV-B_{nd} model fares favorably out-of-sample in a manner consistent with its strong performance in the simulation environment.

6 Conclusion

We introduce a family of tractable no-dominance term structure models where bond prices are analytically available and very nearly arbitrage-free for essentially any specification of the time-series dynamics. Our results based on some special cases show how this new class of models can do a reasonable job, relative to existing models, in capturing the dynamics of conditional bond Sharpe ratios before and after 2008, when yields reach the lower bound.

Variation in the volatility term structure remains a challenge for existing models. Our family of no-dominance models is large and permits flexible specifications of the dynamic interactions between yield and macro variables, including time-varying volatility and the interaction with the lower bound. This should allow future research to revisit several results involving the trade-off between the risk premium and yield volatility faced by investors, the influence of conventional and unconventional policy actions on this trade-off (including quantitative easing and forward guidance), and the correlations among international term structures (when far from or near to their respective lower bounds).

References

Adrian, T., Crump, R. K., Moench, E., 2013. Pricing the term structure with linear regressions. Journal of Financial Economics 110, 110–138.

Ahn, D., Dittmar, R., Gallant, A., 2002. Quadratic term structure models: Theory and evidence. Review of Financial Studies 15, 243–288.

Andreasen, M. M., Christensen, B. J., 2015. The SR approach: A new estimation procedure for non-linear and non-Gaussian dynamic term structure models. Journal of Econometrics 184, 420 – 451.

Banerjee, S., Graveline, J. J., 2013. The cost of short-selling liquid securities. Journal of Finance 68, 637–664.

Bansal, R., Tauchen, G., Zhou, H., 2004. Regime shifts, risk premiums in the term structure, and the business cycle. Journal of Business & Economic Statistics 22, 396–409.

Bauer, M., Rudebusch, G., 2016. Monetary policy expectations at the zero lower bound. Journal of Money, Credit and Banking 48, 1439–1465.

Bauer, M. D., Rudebusch, G. D., 2020. Interest rates under falling stars. American Economic Review 110, 1316–54.

Bjork, T., Christensen, B. J., 1999. Interest rate dynamics and consistent forward rates curves. Mathematical Finance 9, 323–348.

Black, F., 1995. Interest rates as options. Journal of Finance 50, 1371–1376.

Campbell, J., Shiller, R., 1991. Yield spreads and interest rate movements: A bird's eye view. Review of Economic Studies 58, 495–514.

Christensen, J. H., Diebold, F. X., Rudebusch, G. D., 2011. The affine arbitrage-free class of nelson–siegel term structure models. Journal of Econometrics 164, 4–20.

Christensen, J. H. E., Rudebusch, G. D., 2014. Estimating Shadow-Rate Term Structure Models with Near-Zero Yields. Journal of Financial Econometrics 13, 226–259.

Cieslak, A., Povala, P., 2016. Information in the term structure of yield curve volatility. The Journal of Finance 71, 1393–1436.

Collin-Dufresne, P., Goldstein, R. S., 2002. Do bonds span the fixed income markets? Theory and evidence for 'unspanned' stochastic volatility. Journal of Finance 57, 1685–1730.

Coroneo, L., Nyholm, K., Vidova-Koleva, R., 2011. How arbitrage-free is the Nelson-Siegel model? Journal of Empirical Finance 18, 393–407.

Creal, D. D., Wu, J. C., 2015. Estimation of affine term structure models with spanned or unspanned stochastic volatility. Journal of Econometrics 185, 60 – 81.

Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943–1978.

Dai, Q., Singleton, K., 2002. Expectations puzzles, time-varying risk premia, and affine models of the term structure. Journal of Financial Economics 63, 415–441.

Diebold, F., Li, C., 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 130, 337–364.

Diebold, F. X., Rudebusch, G. D., 2012. Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach. Princeton University Press.

Diez de los Rios, A., 2015. A new linear estimator for Gaussian dynamic term structure models. Journal of Business & Economic Statistics 33, 282–295.

Duffee, G., 2002. Term premia and interest rate forecasts in affine models. Journal of Finance 57, 405–443.

Duffee, G., 2010. Sharpe ratios in term structure models. Working paper, Johns Hopkins University.

Duffee, G., 2011. Information in (and not in) the term structure. Review of Financial Studies 24, 2895–2934.

Duffie, D., 1996. Special repo rates. Journal of Finance 51, 493–526.

Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica 50, 987–1008.

Engle, R., Roussellet, G., Siriwardane, E., 2017. Scenario generation for long run interest rate risk assessment. Journal of Econometrics 201, 333 – 347.

Feunou, B., Fontaine, J.-S., 2014. Non-Markov Gaussian term structure models: The case of inflation. Review of Finance 18, 1953–2001.

Filipovic, D., 1999. A note on the Nelson-Siegel family. Mathematical Finance 9, 349–359.

Filipovic, D., Larsson, M., Trolle, A. B., 2017. Linear-rational term structure models. The Journal of Finance 72, 655–704.

Goliński, A., Zaffaroni, P., 2016. Long memory affine term structure models. Journal of Econometrics 191, 33–56.

Gurkanyak, R., Sack, B., Wright, J., 2007. The U.S. Treasury yield curve: 1961 to the present. Journal of Monetary Economics 54, 2291–2304.

Hamilton, J., Wu, J., 2011. Identification and estimation of affine term structure models. Working Paper, University of California, San Diego.

Jacobs, K., Karoui, L., 2009. Conditional volatility in affine term-structure models: Evidence from Treasury and swap markets. Journal of Financial Economics 91, 288–318.

Joslin, S., 2018. Can unspanned stochastic volatility models explain the cross section of bond volatilities? Management Science 64, 1707–1726.

Joslin, S., Le, A., 2021. Interest rate volatility and no-arbitrage affine term structure models. Management Science (forthcoming).

Joslin, S., Le, A., Singleton, K., 2013. Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. Journal of Financial Economics 109, 604–622.

Joslin, S., Priebsch, M., Singleton, K. J., 2014. Risk premiums in dynamic term structure models with unspanned macro risks. The Journal of Finance 69, 1197–1233.

Joslin, S., Singleton, K., Zhu, H., 2011. A new perspective on Gaussian dynamic term structure models. Review of Financial Studies 24, 926–970.

Kim, D., Priebsch, M. A., 2013. The U.S. yield curve at the zero lower bound: Gaussian shadow-rate vs. affine-Gaussian term structure models. Working paper, Federal Reserve Bank of New York.

Kim, D. H., Singleton, K. J., 2012. Term structure models and the zero bound: An empirical investigation of Japanese yields. Journal of Econometrics 170, 32–49.

Kozicki, S., Tinsley, P. A., 2001. Shifting endpoints in the term structure of interest rates. Journal of monetary Economics 47, 613–652.

Krippner, L., 2011. Modifying Gaussian term structure models when interest rates are near the zero lower bound. Working Paper, Reserve Bank of New Zealand.

Krippner, L., 2013. A theoretical foundation for the Nelson-Siegel class of yield curve models. Journal of Applied Econometrics 30, 97–118.

Krishnamurthy, A., 2002. The bond/old-bond spread. Journal of Financial Economics 66, 463–506.

Leippold, M., Wu, L., 2003. Design and estimation of quadratic term structure models. European Finance Review 7, 47–73.

Levy, H., 1992. Stochastic dominance and expected utility: Survey and analysis. Management Science 38, 555–593.

Li, H., Zhao, F., 2006. Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. Journal of Finance 61, 341–378.

Monfort, A., Pegoraro, F., Renne, J.-P., Roussellet, G., 2017. Staying at zero with affine processes: An application to term structure modelling. Journal of Econometrics 201, 348 – 366.

Nelson, C., Siegel, A., 1987. Parsimonious modelling of yield curves. Journal of Business 60, 473–489.

Noureldin, D., Shephard, N., Sheppard, K., 2011. Multivariate high-frequency-based volatility (HEAVY) models. Journal of Applied Econometrics 27, 907–933.

Priebsch, M., 2013. Computing arbitrage-free yields in multi-factor Gaussian shadow-rate term structure models. Tech. rep., Finance and Economics Discussion Series, The Federal Reserve Board.

Realdon, M., 2006. Quadratic term structure models in discrete time. Finance Research Letters 3, 277–289.

Rothschild, M., Stiglitz, J., 1970. Increasing risk: I. A definition. Journal of Economic Theory 2, 225–243.

Swanson, E. T., Williams, J. C., 2014. Measuring the effect of the zero lower bound on medium- and longer-term interest rates. The American Economic Review 104, 3154–3185.

Vayanos, D., Weill, P.-O., 2008. A search-based theory of the on-the-run phenomenon. Journal of Finance 63, 1361–1398.

Wu, J. C., Xia, F. D., 2016. Measuring the macroeconomic impact of monetary policy at the zero lower bound. Journal of Money, Credit and Banking 48, 253–291.

A Appendix

A.1 No-Dominance Term Structure Models

In the following, we study dominant and arbitrage trading strategies between bonds. For this purpose, note that, starting with $P_0(\cdot) \equiv 1$ and expanding the recursion (1), we get

$$P_n(X_t) = \exp(-\sum_{i=0}^{n-1} m(g^{\circ i}(X_t))). \tag{A.1}$$

Equation (3) follows from the definition of the n-period yield and forward rate, $y_{n,t} \equiv -\log(P_n(X_t))/n$ and $f_{n,t} \equiv (n+1)y_{n+1,t} - ny_{n,t}$, respectively. In the following results, we rely on Assumption 2, which is essential to clarify the set of admissible time series dynamics.

Assumption 2. *The time series dynamics of X_t admit $\underline{X}$ as support.*

Assumption 2 is analogous to the requirement in NA models that the historical and risk-neutral measures be equivalent measures. This is a mild requirement, since this affords researchers the flexibility to consider rich dynamic and distributional assumptions. First, Theorem 1 checks dominant bond trading strategies; portfolios of bonds with strictly positive payoffs.

Theorem 1 (No Dominant Strategies)**.** *Assumptions 1-2 guarantee the absence of strictly dominant trading strategies in the bond markets.*

Proof. Let w_n denote the amount (in face value) invested in each n-period bond. Suppose that this portfolio guarantees positive payoffs: $\sum_n w_n P_{n-1}(X_{t+1}) > 0 \; \forall X_{t+1} \in \underline{\mathrm{X}}$. From the pricing recursions in Equation (1), the price of this portfolio is given by

$$\sum_n w_n P_n(X_t) = \exp(-m(X_t)) \times \sum_n w_n P_{n-1}(g(X_t)). \tag{A.2}$$

Since $g(X_t) \in \underline{\mathrm{X}}$, and since $\sum_n w_n P_{n-1}(X_{t+1}) > 0$ for all $X_{t+1} \in \underline{\mathrm{X}}$, it follows that the price of this portfolio is strictly positive. Thus, a dominant trading strategy does not exist. □

Next, we consider portfolios that have a zero payoff with strictly positive probability, but positive payoffs otherwise. The NA condition requires that the price of this portfolio be strictly positive. Theorem 2 establishes that portfolios like this one cannot admit strictly negative prices in our framework.

Theorem 2 (Non-Negative Payoffs)**.** *Assumptions 1-2 ensure that bond portfolios with strictly non-negative payoffs cannot admit strictly negative prices.*

Proof. Let w_n denote the amount (in face value) invested in each n-period bond. Consider a portfolio with strictly non-negative payoffs: $\sum_n w_n P_{n-1}(X_{t+1}) \geq 0 \; \forall X_{t+1} \in \underline{\mathrm{X}}$. From the pricing recursion in Equation (1), the price of this portfolio is given by

$$\sum_n w_n P_n(X_t) = \exp(-m(X_t)) \times \sum_n w_n P_{n-1}(g(X_t)). \tag{A.3}$$

The price of this portfolio cannot be negative, for it requires $\sum_n w_n P_{n-1}(g(X_t)) < 0$, but this would contradict $g(X_t) \in \underline{\mathrm{X}}$ and $\sum_n w_n P_{n-1}(X_{t+1}) \geq 0$ for all $X_{t+1} \in \underline{\mathrm{X}}$. □

Figure A.1 illustrates Theorem 2 and clarifies that what is setting NA and ND apart is the set of zero-cost portfolios with zero or positive payoffs (both with positive probability). For portfolios with strictly non-negative payoffs, the absence of arbitrage restricts prices to be on the positive half of the real line, excluding the origin. Assumption 1 allows for bond prices on the positive half of the real line, including the origin. The difference reduces to one point on the real line, the origin, graphically illustrated by Figure A.1.

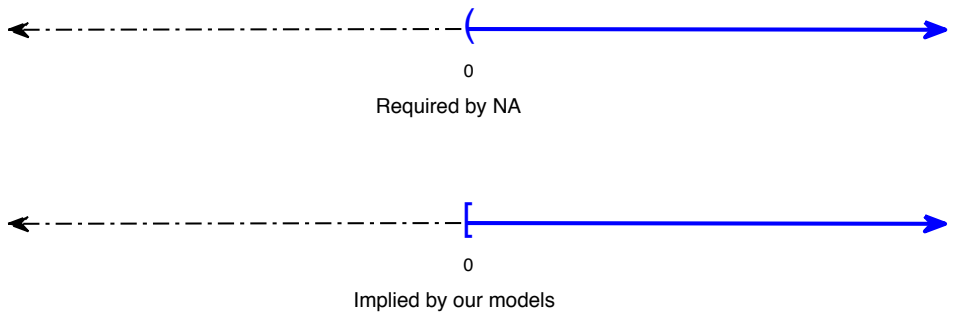


Figure A.1: Prices of portfolios with strictly non-negative payoffs

Figure A.1 suggests that the distance between our models and the no-arbitrage paradigm is small. Does this difference matter? Note that Figure A.1 also tells us that any remaining arbitrage opportunities must involve a self-financing portfolio with strictly non-negative payoffs in the future and with strictly zero set-up costs today. But a self-financing portfolio must involve short-selling bonds and, since short positions imply non-zero costs to set up and to maintain, the specific portfolio in Figure A.1 may not represent an arbitrage opportunity in a practical sense for sophisticated long-short investors. The presence of small transaction costs preventing self-financing strategies is economically plausible.

Theorem 3 (Transaction Costs)**.** *Assumptions 1-2 rule out all arbitrage opportunities in bond prices if transaction costs for short positions are positive.*

Proof. Consider again a portfolio in which the amount (in face value) invested in each n-period bond is given by w_n. To account for transaction costs, let's assume that the set-up cost at time t of the portfolio is C_0 and, to realize the cash flows at time $t+1$, the transaction cost is given by C_1. The existence of transaction costs means that at least one of C_0 and C_1 is strictly positive. The one-period-ahead cash flows net of transaction costs are given by

$$CF(X_{t+1}) = \sum_n w_n P_{n-1}(X_{t+1}) - C_1.$$

Recall that the absence of arbitrage is equivalent to requiring that any portfolio with non-negative payoffs must command a positive price. Thus, the key question here is the following: if $CF(X_{t+1}) \geq 0$ for every $X_{t+1} \in \underline{X}$, then can we show that the price of the portfolio must be strictly positive, net of transaction costs? The price of this portfolios is given by

$$\begin{aligned} Price(X_t) &= \sum_n w_n P_n(X_t) + C_0 \\ &= \exp(-m(X_t)) \sum_n w_n P_{n-1}(g(X_t)) + C_0, \\ &= \exp(-m(X_t)) \times (CF(g(X_t)) + C_1) + C_0, \\ &= \exp(-m(X_t)) \times CF(g(X_t)) + \exp(-m(X))C_1 + C_0, \end{aligned} \tag{A.4}$$

where the second line follows from Equation (1) and the third line follows from the definition of the cash flows $CF(X_{t+1})$. The first term on the right-hand side of (A.4) is non-negative, since $g(X_t) \in \underline{X} \Rightarrow CF(X_t) \geq 0$.

Additionally, the last two terms must add up to a strictly positive number, since C_0 and C_1 cannot be jointly zero. Thus, $Price(X_t) > 0$ as needed. $\square$

Online Appendix

I Special Cases

Nelson-Siegel Let Δ denote the discrete time interval and consider a three-factor model $N = 3$ with the following one-period interest rate:

$$m(X_t) = \Delta(1, 1, 0)' X_t$$

and the $g(\cdot)$ function

$$g(X_t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 1-a \\ 0 & 0 & a \end{pmatrix} X_t.$$

Using the pricing equations, the per annum yield for a bond with maturity $m = n\Delta$ is given by:

$$y_{n,t}/\Delta = X_{1,t} + (X_{2,t} + X_{3,t})\frac{1-a^n}{1-a}/n - a^{n-1}X_{3,t}. \tag{B.5}$$

Then, using notations of Nelson and Siegel (1987): $a = 1 - \Delta/\tau$ where $1/\tau$ can be interpreted as modulating the frequency of the factors, we can use the L'Hôpital's rule to show that as $\Delta \to 0$,

$$a^{n-1} \to e^{-m/\tau} \text{ and } \frac{1-a^n}{1-a}/n \to (1 - e^{-m/\tau})/(m/\tau),$$

and, therefore, that the yield in (B.5) approaches:

$$X_{1,t} + (X_{2,t} + X_{3,t})(1 - e^{-m/\tau})/(m/\tau) - e^{-m/\tau}X_{3,t}.$$

This last expression is identical to the model of Nelson and Siegel (1987) with $X_{1,t}$, $X_{2,t}$, and $X_{3,t}$ corresponding respectively to their original parameters of β_0, β_1, and β_2.

Linear models Suppose $X_t \in \mathbb{R}^N$. The following natural specification leads to affine Gaussian term structure models:

$$m(X_t) = \delta_0 + \delta_1' X_t \tag{B.6}$$
$$g(X_t) = KX_t, \tag{B.7}$$

where δ_0 is a scalar, δ_1 is an $N \times 1$ vector and K is an $N \times N$ matrix. From Equation (A.1), yields are linear:

$$y_{n,t} = \delta_0 + (B_n/n)X_t, \tag{B.8}$$

with B_n given by the recursion $B_n = B_{n-1}K + \delta_1'$. For comparison, the $A_0(N)$ Gaussian dynamic term structure model (e.g., Dai and Singleton, 2000 and Duffee, 2002) has a linear short rate equation and risk-neutral dynamics given by

$$\begin{aligned} r_t &= \delta_0 + \delta_1' X_t \\ X_{t+1} &= K_0 + K_1^{\mathbb{Q}} X_t + \epsilon_{t+1}, \end{aligned} \tag{B.9}$$

where $\epsilon_{t+1} \sim N(0, \Sigma)$. The solution for yields in that standard case is given by

$$y_{n,t} = A_n/n + (B_n/n)X_t, \tag{B.10}$$

with coefficients given by

$$B_n = B_{n-1}K_1^{\mathbb{Q}} + \delta_1', \tag{B.11}$$

$$A_n = A_{n-1} + \delta_0 - \frac{1}{2}B_{n-1}\Sigma B_{n-1}'. \tag{B.12}$$

Clearly, the short rate r_t and the loadings B_n are identical between these models. The intercept terms A_n for $n > 1$ are different only because of the convexity correction $B_{n-1}\Sigma B_{n-1}'$. This Jensen term is negligible in a typical application (see Figure B.2).

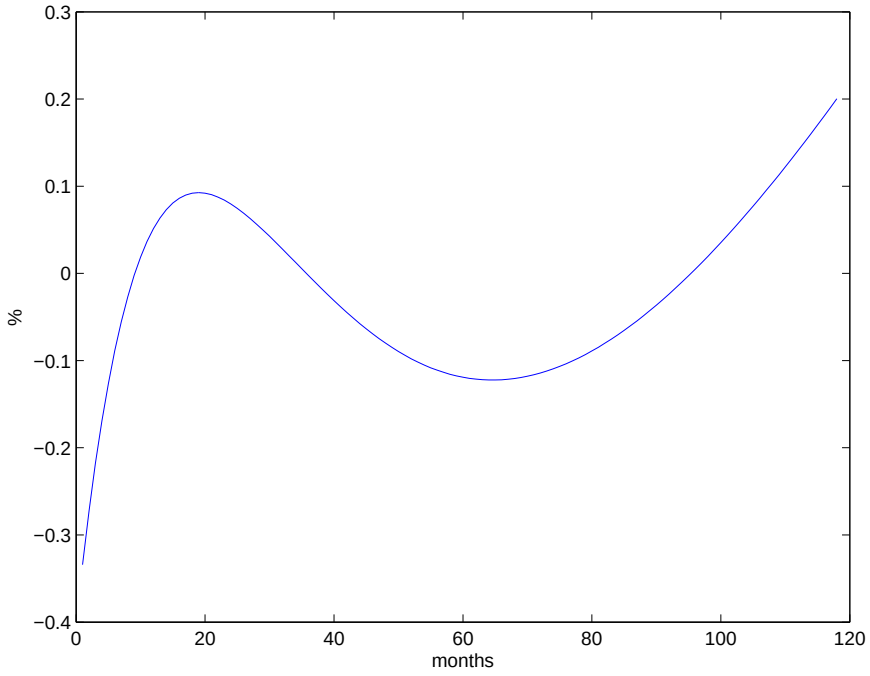


Figure B.2: The Jensen term ($\frac{1}{2}B_{n-1}\Sigma B_{n-1}'$) is negligible in Gaussian models. Results are based on parameter estimates from the canonical representation in Joslin, Singleton, and Zhu (2011).

Quadratic models The following choice generates linear-quadratic TTSMs:

$$\begin{aligned} m(X_t) &= \delta_0 + \delta_1' X_t + X_t'\delta_2 X_t \\ g(X_t) &= K X_t, \end{aligned} \tag{B.13}$$

where δ_0 is a scalar, δ_1 is a $N \times 1$ vector and δ_2 is a $N \times N$ matrix. Using our pricing equation, yields are given by

$$y_{n,t} = \delta_0 + (B_n/n)X_t + X_t'(C_n/n)X_t, \tag{B.14}$$

where the linear and quadratic coefficients B_n and C_n are given by

$$\begin{aligned} C_n =& K' C_{n-1} K + \delta_2 \\ B_n =& B_{n-1} K + \delta_1'. \end{aligned} \tag{B.15}$$

Compare this with the affine-quadratic model developed by Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2003) (see Realdon (2006) for discrete-time treatment) where the short rate equation is quadratic:

$$r_t = \delta_0 + \delta_1' X_t + X_t' \delta_2 X_t,$$

and with the risk-neutral dynamics as in (B.9). The solution for yields in this case is given by

$$y_{n,t} = A_n/n + (B_n/n) X_t + X_t' (C_n/n) X_t, \tag{B.16}$$

where the loadings A_n, B_n and C_n are given by the following recursions:

$$\begin{aligned} C_n =& K_1^{\mathbb{Q}\prime} C_{n-1} \Omega_{n-1} K_1^{\mathbb{Q}} + \delta_2, \\ B_n =& B_{n-1} \Omega_{n-1} K_1^{\mathbb{Q}} + \delta_1', \\ A_n =& A_{n-1} + \delta_0 - \frac{1}{2} log|\Omega_{n-1}| - \frac{1}{2} B_{n-1} \Omega_{n-1} \Sigma B_{n-1}, \end{aligned} \tag{B.17}$$

with $\Omega_{n-1} \equiv (I_N - 2\Sigma C_{n-1})^{-1}$. Comparing loadings in (B.15) and (B.17) reveals two differences. First, the term $B_{n-1}\Omega_{n-1}\Sigma B_{n-1}$ reflects a convexity adjustment. Second, the matrix Ω_{n-1} may introduce a wedge between loadings if the quadratic coefficient δ_2 is "large."

II Estimation and Implementation Strategy

Portfolios of Shadow Forwards as Risk Factors The ND pricing equation (7):

$$f_{n,t} = \theta \, w \left(\frac{\delta_0 + \delta_1' K^n X_t}{\theta} \right)$$

is strictly increasing and has a well-defined inverse $f^{-1}(\cdot)$. Applying the inverse $f^{-1}(\cdot)$ to both sides, we can write:

$$s_{n,t} \equiv \theta \, w^{-1} \left(\frac{f_{n,t}}{\theta} \right) = \delta_0 + \delta_1' K^n X_t. \tag{B.18}$$

We refer to the left-hand side $s_{n,t}$ of Equation (B.18) as the shadow forward implied in the Black style model. Shadow forwards are linear in the states, and thus can turn negative, just as the shadow short rate. In the case $n = 0$, the shadow forward is the same as the shadow rate $s_{0,t} = s_t$.

Given a value of the scaling parameter θ, the portfolios of shadow forwards are directly observable within our model from Equation (B.18). Now consider J shadow forwards $s_{n,t}$ with maturities $n = n_1, n_2, \ldots, n_J$ stacked into a $J \times 1$ vector $\bar{s}_t$:

$$\bar{s}_t = A_X + B_X X_t \tag{B.19}$$

where A_X and B_X stack the intercepts and loadings of Equation (B.18), respectively. Equation (B.19) is linear. Therefore, we can apply the insight of JSZ and rotate the latent state vector X_t into linear combinations of shadow forwards. Specifically, we construct N portfolios of shadow forwards $\mathcal{P}_t = W \bar{s}_t$ for

a given $N \times J$ matrix W. We can recover the latent state X_t as a linear function of the portfolio $\mathcal{P}_t$,

$$X_t = (WB_X)^{-1}(\mathcal{P}_t - WA_X),$$

which can be substituted into Equation (B.19) to obtain a pricing equation for the shadow forwards in terms of $\mathcal{P}_t$:

$$\bar{s}_t = A_{\mathcal{P}} + B_{\mathcal{P}}\mathcal{P}_t, \tag{B.20}$$

where $B_{\mathcal{P}} = B_X(WB_X)^{-1}$ and $A_{\mathcal{P}} = A_X - B_{\mathcal{P}}(WA_X)$. Equipped with the shadow forwards, we can evaluate the actual forwards simply by $f_{n,t} = \theta\, w\left(\frac{s_{n,t}}{\theta}\right)$, which only involves the parameters θ, δ_0, δ_1 and K governing bond pricing. We also impose additional requirements to ensure that near-zero values of the shadow rate s_t, between 5 and 50 bps, correspond to values of the short rate near the lower bound. In particular, we require: (i) $\theta\, w\left(\frac{0-\sigma_s}{\theta}\right) \geq 0.0005$; and (ii) $\theta\, w\left(\frac{0+\sigma_s}{\theta}\right) \leq 0.0050$ where σ_s is the volatility of the shadow short rate s_t.

Identification The ND that we propose in Section 2.1 has the parameters θ, δ_0, δ_1 and K governing bond pricing in addition to the parameters K_0, K_1, Σ_0, a, and b that are responsible for the time series dynamics. Not all of the parameters are econometrically identified, because the state variables are latent. We adopt the JSZ canonical form whereby all identification assumptions are implemented on the pricing side. Specifically, we assume that $\delta_1 = \iota$ is a vector of ones and K has an ordered Jordan form. The JSZ canonical form leaves the time series dynamics completely unconstrained. Therefore, this identification strategy is applicable to the rich set of time series dynamics satisfying Assumption 2. This identification approach will also play an important role in constructing the likelihood and analytically concentrating some parameters in the estimation. Based on the insights of JSZ, we anticipate that estimation of the $\mathcal{P}$-representation will be robust.

Conditional Likelihood It is easy to see that $\mathcal{P}_t$ will inherit the VAR(1) structure as well as the scalar BEKK volatility specification from the X-dynamics, since the correspondence between X_t and $\mathcal{P}_t$ is linear. We have:

$$\mathcal{P}_{t+1} = K_{0\mathcal{P}} + K_{1\mathcal{P}}\mathcal{P}_t + \sqrt{\Sigma_{t,\mathcal{P}}}\varepsilon_{t+1}, \tag{B.21}$$

$$\Sigma_{t,\mathcal{P}} = \Sigma_{0,\mathcal{P}} + a\Sigma_{t,\mathcal{P}} + b\Delta\mathcal{P}_t\Delta\mathcal{P}_t'. \tag{B.22}$$

Therefore, the primitive parameters of our model governing the dynamics of $\mathcal{P}$ are $K_{0,\mathcal{P}}$, $K_{1,\mathcal{P}}$, $\Sigma_{0,\mathcal{P}}$, a, and b. Importantly, due to the JSZ normalization, these parameters are unconstrained. The parameters θ, δ_0 and K governing the pricing equation remain unchanged.

We estimate our models by maximizing the log likelihood of the observed forwards.[8] That is, we need to compute for each t:

$$\mathbb{P}(f_t^o|\mathcal{I}_{t-1}),$$

where f_t^o denotes the $J \times 1$ vector $(f_{n_1,t}^o, f_{n_2,t}^o, \ldots, f_{n_J,t}^o)'$ of observed forwards and $\mathcal{I}_t$ denotes the information set generated by f_t^o up to time t. The superscript ^{o} differentiates observed quantities from their theoretical constructs. The overall likelihood is simply obtained by $\prod_t \mathbb{P}(f_t^o|\mathcal{I}_{t-1})$. We construct this likelihood as follows. We make the assumption that the N portfolios of shadow forwards $\mathcal{P}_t^o$ are priced without error and identical to their model counterparts at each point in time: $\mathcal{P}_t^o = W\bar{s}_t^o \equiv W\bar{s}_t = \mathcal{P}_t$. Next, we assume

[8] Recent advances provide alternative estimation methods. See, for example, Joslin, Singleton, and Zhu (2011), Joslin, Le, and Singleton (2013), Hamilton and Wu (2011), Adrian, Crump, and Moench (2013), and Diez de los Rios (2015).

that $J-N$ combinations of forwards $W_e f_t$ are priced with i.i.d. errors:

$$W_e(f_t^o - f_t) \sim N(0, \sigma_e^2 I_{J-N}), \tag{B.23}$$

given some $(J-N)\times J$ loading matrix W_e. These two assumptions are similar to case **C** in JSZ where the likelihood of coupon bonds is used to estimate the model using N combinations of zero-coupon yields measured without errors, which are a non-linear transformation of the coupon bond yields. The analogy is that here the shadow forwards are a non-linear transformation of the forwards.

Combining these two measurement assumptions, we can write the one-step-ahead conditional likelihood of forwards f_t^o:

$$\mathbb{P}(f_t^o|\mathcal{I}_{t-1}) = \mathbb{P}\left(W_e f_t^o|\mathcal{P}_t, \mathcal{I}_{t-1}\right) \times \mathbb{P}\left(\mathcal{P}_t|\mathcal{I}_{t-1}\right) \times \left|\frac{\partial h(f_t^o)}{\partial f_t^o}\right|, \tag{B.24}$$

which have three components on the right-hand side. Except for the SV-B_{nd} model, it is straightforward to derive the model-implied likelihood of the data for the more general case in which all forwards are observed with errors, using the Kalman filter.

The first component $\mathbb{P}\left(W_e f_t^o|\mathcal{P}_t, \mathcal{I}_{t-1}\right)$ captures the cross-sectional fit of the model: its ability to explain the observed forwards $W_e f_t^o$ based on N portfolios of shadow forwards $\mathcal{P}_t$ observed contemporaneously. This component can be computed easily using the distribution of pricing errors assumed in Equation (B.23). The second component $\mathbb{P}\left(\mathcal{P}_t|\mathcal{I}_{t-1}\right)$ captures the time series fit of the model: the predictive density of observing $\mathcal{P}_t$ given the information set one period earlier. This component can also be derived in a straightforward manner, since $\mathcal{P}_t$ is measured without errors and governed by the conditional Gaussian VAR(1) dynamics in Equations (B.21-B.22). The final term is a Jacobian adjustment to account for the non-linear dependence between forward rates and their shadow counterparts:

$$h(f_t^o) = \begin{pmatrix} W_e f_t^o \\ W\theta\, w^{-1}(f_t^o/\theta) \end{pmatrix}.$$

To avoid a singular likelihood, the choice of the matrices W and W_e must be such that the Jacobian adjustment term is non-zero. For simplicity, we choose W as the loadings on the first N PCs of f_t^o and W_e the remaining $J-N$ PCs of f_t^o.

Analytical Concentration of Parameters The maximum likelihood estimates of $K_{0\mathcal{P}}$ and $K_{1\mathcal{P}}$ can be derived analytically. This is possible because these conditional mean parameters do not mix with volatility parameters in the expression for the log likelihood. Specifically, $K_{0\mathcal{P}}$ and $K_{1\mathcal{P}}$ enter the log likelihood via a quadratic form because time series innovations are conditionally Gaussian. Therefore, the first-order derivative of the log likelihood must be linear and the maximum likelihood estimates $K_{0\mathcal{P}}$ and $K_{1\mathcal{P}}$ are given:

$$vec([\hat{K}_{0\mathcal{P}}, \hat{K}_{1\mathcal{P}}]) = E_T[\mathcal{P}_t^a \mathcal{P}_t^{a\prime} \otimes \Sigma_t^{-1}]^{-1} vec(E_T[\Sigma_t^{-1}\mathcal{P}_{t+1}\mathcal{P}_t^{a\prime}]), \tag{B.25}$$

where $E_T[.]$ denotes sample average, $\otimes$ is the Kronecker product, $\mathcal{P}_t^a = (1,\ \mathcal{P}_t')'$ and where the $vec(\cdot)$ converts a matrix into a column vector by stacking its columns (from left to right). This result arises because we use the first difference $\Delta\mathcal{P}_t$ to update volatility in Equation (B.22), instead of the innovations $\mathcal{P}_{t+1} - (K_{0\mathcal{P}} + K_{1\mathcal{P}}\mathcal{P}_t)$, which would be the case in the standard BEKK volatility. This alternative construction ensures that the parameters $K_{0\mathcal{P}}$ and $K_{1\mathcal{P}}$ do not enter the dynamics of Σ_t. Otherwise, the first-order conditions of the log likelihood with respect to $K_{0\mathcal{P}}$ and $K_{1\mathcal{P}}$ would be highly non-linear and closed-form estimates of $K_{0\mathcal{P}}$ and $K_{1\mathcal{P}}$ would likely be much harder to obtain.

III Simulation Exercise

Data generating process We use the linear rational term structure models of Filipovic et al. (2017) to simulate samples of yields data. In particular, we use their LRSQ(3,3) specification with three factors Z_t explaining the cross-section of bond yields and three factors U_t driving unspanned yields volatility. The state 6×1 state vector X_t follows an extended square root process:

$$dX_t = (b - \beta X_t)dt + diag(\sigma)\sqrt{X_t}dB_t,$$

the volatility factors U_t correspond to the last three elements of X_t:

$$U_t = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} X_t,$$

and three cross-sectional factors Z_t correspond to the sum of the first three elements of X_t and the last three elements of X_t:

$$Z_t = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} X_t.$$

Pricing functions The pricing functions involve five parameters: $\alpha, \phi, \psi, \theta, \kappa$. The price of a zero-coupon bond with τ years to maturity is given by:

$$p_t = e^{-\alpha\tau}\frac{\phi + \psi'(\theta + e^{-\tau\kappa}(Z_t - \theta))}{\phi + \psi' Z_t},$$

which is the ratio of two linear functions of Z, hence the name linear rational. The corresponding yield to maturity is given by:

$$y_t = -\frac{log(p_t)}{\tau},$$

and, taking the limit $\tau = 0$, the instantaneous short rate is given by:

$$r_t = \alpha + \frac{\psi'\kappa(Z_t - \theta)}{\phi + \psi' Z_t}.$$

Generating ZLB samples We discretize the dynamics of X_t:

$$X_{t+\Delta} = max(X_t + (b - \beta X_t)\Delta + diag(\sigma\sqrt{\Delta})\sqrt{X_t}\epsilon_{t+\Delta}, 10^{-8}),$$

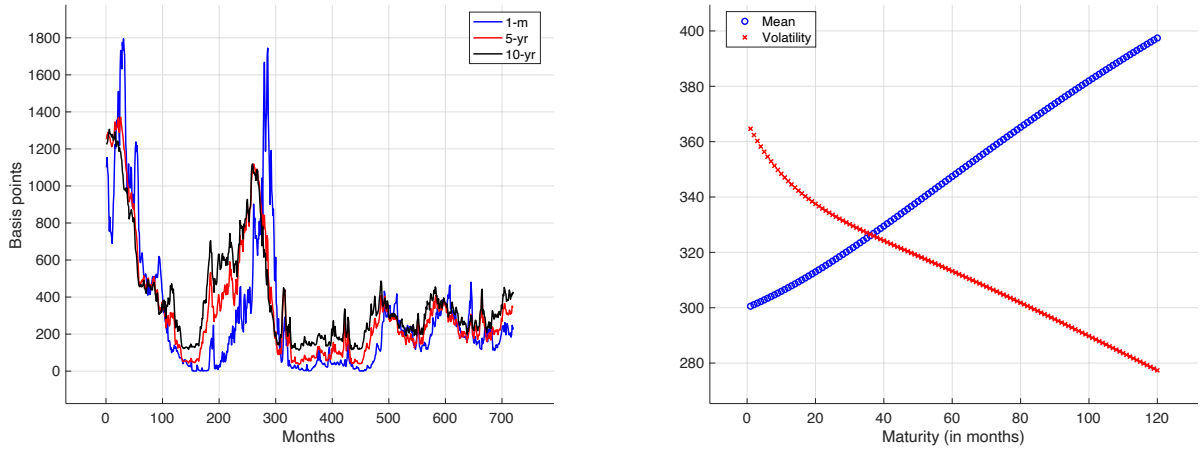
where the frequency is daily $\Delta = 1/365$ and where $\epsilon_{t+\Delta}$ is an i.i.d. standard normal variable. Using these dynamics, we first run a burn-in sample of 10 years and then generate a daily sample of 1000 years. Out of this sample, we randomly choose a ZLB episode, defined as any single day for which the short rate is less than or equal to one basis point. Starting from this day, we go back by 30 years and collect end-of-month yields. This design implies that ZLB episodes are typically found toward the end of the simulated samples. We repeat this process until we obtain 100 different simulated samples, each of which featuring the ZLB behavior of interest rates. If needed, we generate another daily sample of 1000 years with different simulation seeds until we reach 100 different simulated samples.

Parameter values We use parameter estimates and normalized values provided in Filipovic et al. (2017)$\alpha = 0.0566$, $\phi = 1$, $\psi' = (1, 1, 1)'$, $\theta' = (0.2985, 0.5738, 0.1601)'$ and

$$\kappa = \begin{bmatrix} 0.1895 & 0 & 0 \\ -0.2460 & 0.1280 & 0 \\ 0 & -0.1846 & 0.6616 \end{bmatrix}, \sigma = \begin{bmatrix} 0.3177 \\ 0.2676 \\ 0.0509 \\ 0.7920 \\ 0.9135 \\ 0.3585 \end{bmatrix}, b \times 1000 = \begin{bmatrix} 44.5325 \\ 0.0076 \\ -0.0222 \\ 12.0333 \\ 0.0078 \\ 0.0209 \end{bmatrix},$$

$$\beta = \begin{bmatrix} 0.1895 & 0 & 0 & 0 & 0 & 0 \\ -0.2460 & 0.1280 & 0 & 0 & 0 & 0 \\ 0 & -0.1846 & 0.6616 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1895 & 0 & 0 \\ 0 & 0 & 0 & -0.2460 & 0.1280 & 0 \\ 0 & 0 & 0 & 0 & -0.1846 & 0.6616 \end{bmatrix}.$$

Example Panel (a) of Figure B.3 shows the yields with maturity of 1 month, 5 years, and 10 years in one simulated sample with 60 years of data. This provides a "big picture" view. The simulated yields also exhibit significant variation of yield volatilities over time and can range as high as 10 percent. Each of the simulation samples lasts 30 years and features lower-bound episodes by design of the simulation. Panel (b) of Figure B.3 shows the average term structures of yields and volatilities across simulations. The average yield curve is upward sloping and the average volatility curve is downward sloping—consistent with stylized facts in the U.S. data. Appendix III provides a more detailed description of the linear-rational model, the estimates that we use, and the procedure to simulate yield samples from the model.



(a) One sample of simulated yields

(b) Average means and volatilities of simulated yields

Figure B.3: Simulated Yields

Panel (a): the time series of yields with 1 month, 5 years and 10 years to maturity in the first simulated sample. Panel (b): the average yields and the average yield volatilities in the first simulated sample.

IV Implementation of Wu and Xia (2016)

Wu and Xia (2016) implement the short rate equation $r_t = \max(\underline{r}, s_t)$. For convenient estimation, it is preferable that the short rate function be invertible. To circumvent this issue, we start with their NA model generated at a daily frequency but that we implement with a monthly sampling frequency. In this case, the one-month short rate can be inverted for all values of the shadow rate s_t. Assume that the daily risk-neutral dynamics are given by:

$$X_{t+1} = \mu^Q + \rho^Q X_t + \sigma \epsilon^Q_{t+1},$$

with $\epsilon^Q_{t+1} \sim N(0, I)$. As usual in this case, the shadow rate $s_t = \delta_0 + \delta_1' X_t$ also defines the short rate $r_t = max(\underline{r}, s_t)$ where $\underline{r}$ is the lower bound. We choose the standard normalization $\mu^Q = 0$ and ρ^Q with Jordan form. This implies the monthly dynamics

$$X_{t+1} = K_1^Q X_t + \Sigma \epsilon^Q_{t+1},$$

and forward rate starting at time $n-1$ and maturing at time n given in closed form

$$f_{n-1,n,t} = \underline{r} + S_n g\left(\frac{A_n + B_n' X_t - \underline{r}}{S_n}\right),$$

where $g(z) = z\Phi(z) - \phi(z)$. Note that the forward rate pricing function can be easily inverted for the purpose of estimation:

$$\underline{r} + S_n g^{-1}\left(\frac{f_{n,n+1,t} - \underline{r}}{S_n}\right) = A_n + B_n' X_t,$$

since $g(\cdot)$ is monotone and its inverse function exists. The coefficients S_n, A_n and B_n are given by

$$S_n = \sum_{i=(n-1)\times 30}^{n\times 30-1} \sigma_i, \quad A_n = \sum_{i=(n-1)\times 30}^{n\times 30-1} a_i, \quad B_n = \sum_{i=(n-1)\times 30}^{n\times 30-1} b_i, \tag{B.26}$$

with a_i, b_i and σ_i given by:

$$\begin{aligned} a_n =& \delta_0 - \frac{1}{2}\delta_1' \left(\sum_{j=0}^{n-1} (\rho^Q)^j\right) \sigma \left(\sum_{j=0}^{n-1} (\rho^Q)^j\right)' \delta_1 \\ b_n' =& \delta_1' (\rho^Q)^n, \quad \sigma_n^2 = \sum_{j=0}^{n-1} \delta_1' (\rho^Q)^j \sigma (\rho^{Q\prime})^j \delta_1, \end{aligned}$$

and where

$$\begin{aligned} \rho^Q =& K_1^{Q\,1/30} \quad , \delta_1' = \iota' \left(\sum_{n=0}^{29} \rho^{Q^n}\right)^{-1}, \\ vec(\sigma) =& \left(\sum_{i=0}^{29} \rho^{Q^i} \otimes \rho^{Q^i}\right)^{-1} vec(\Sigma), \\ \delta_0 =& \frac{1}{2}\delta_1' \left[\sum_{n=0}^{29} \left(\sum_{j=0}^{n-1} (\rho^Q)^j\right) \Sigma_d \left(\sum_{j=0}^{n-1} (\rho^Q)^j\right)'\right] \delta_1 + r_\infty^Q, \end{aligned}$$

so that, at the monthly frequency, the model is fully characterized by the parameter set $\{K_0, K_1, \Sigma, r_\infty^Q, K_1^Q\}$.

V Additional Simulation Results

Volatility forecasts Table B.1 reports the performance volatility forecasts. First, the volatility comparison between the SV-B_{nd} model and other models is consistent with Sharpe ratio results. The SV-B_{nd} model consistently delivers the best volatility forecasts over the full sample, for all combinations of horizons and bond maturities. The results are not surprising, since the SV-B_{nd} model has a built-in stochastic volatility construction. This model is able to bring down the RMedSE substantially relative to the G_{na} benchmark, by as much 50-70 percent in many cases. The magnitude of these gains underlines the importance of time-varying volatilities.

Second, in contrast with the Sharpe ratio results, there is no clear winner between the volatility forecasts from the ND B_{nd} and the NA B_{na} in the full sample results. If anything, the B_{nd} model seems to do a better job at the one-year forecasting horizon. There is no clear winner close to the lower bound either. In fact, it is remarkable that the volatility forecasts from the B_{nd} and B_{na} deliver similar performances relative to SV-B_{nd} forecasts, in lower-bound samples, and even better performances for a good number of horizon and bond maturity combinations.

This is remarkable because both models assume constant variances in the state dynamics. The volatility forecasts around the lower bound clearly show the effect of the volatility compression (see e.g., Christensen and Rudebusch 2014; Kim and Priebsch 2013) arising because of the convexity of the (non-linear) transformation between forwards and shadow forwards. Intuitively, as the shadow forwards venture deep into the negative region, the corresponding forwards have little room to "wiggle", thus their conditional volatilities are "compressed" toward zero. On the other hand, when far away from the lower bound, the forwards inherit the constant variance property of the shadow forwards.

Yield forecasts Table B.2 reports each model's performance. First, it is remarkable that the SV-B_{nd} delivers strictly superior full-sample forecasts, and relative to every other model that we consider. The improvements range around 15-25 percent relative to the NA Gaussian and Black models and around 10-20 percent for the B_{nd} model. What is remarkable is that the improvements in the forecasts of yields must be attributed to better forecasts of yield volatilities, which is what distinguishes the SV-B_{nd} model from the others.

Recall from Table B.1 that, across the full sample results, the SV-B_{nd} model captures the volatility dynamics particularly well. In a typical maximum likelihood estimation, it is well known that (i) the variance of forecast errors depends on the variance of the parameter estimator and (ii) that controlling for conditional variance reduces the variance of the parameter estimator (e.g., GLS). Intuitively, ex ante noisier signals have less influence on the estimation. Knowledge of the volatility dynamics may lead to smaller weights in the likelihood function to forecast errors around heightened volatility. In contrast, constant-variance models assign the same weights to forecast errors observed during period of high and low volatility, thereby treating noisy signals and high quality signals equally.

In this light, the ability to match volatility dynamics over the full sample results likely explains the difference between the performances of the B_{nd} and SV-B_{nd} models. By that logic, the differences should be small when these two models produce similar volatility forecast performances. We can check this prediction in the simulations. Indeed, we find the performances of yield forecasts are similar in the lower-bound samples, precisely when Table B.1 shows that the B_{nd} model produces accurate volatility forecasts. The "compression" channel seems to endow the B_{nd} model with as much econometric efficiency gain as the SV-B_{nd} model. In fact, during the lower-bound episodes, the B_{nd} model seems to do a better job at yield forecasting for the shorter bond maturities (1 month and 2 years). Interestingly, these are essentially the categories over which the B_{nd} model provides better volatility forecasts.

		Entire Sample								ZLB Sample							
h	m	μ	σ	T	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	μ	σ	T	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
3m	1m	118.7	106.7	37.9	54.6	1.21	0.30	0.49	0.29*	83.3	96.4	23.2	68.9	1.19	0.15	0.14*	0.15
	2y	86.7	67.4	26.2	31.1	1.18	0.49	0.62	0.33*	62.6	62.0	19.0	36.8	1.14	0.34	0.18*	0.20
	5y	74.2	50.1	21.0	19.9	1.10	0.73	0.73	0.47*	54.8	41.9	15.7	25.2	0.98	0.42	0.42	0.29*
	10y	65.4	41.3	17.0	14.8	0.99	0.58*	0.95	0.62	49.8	32.2	14.3	20.1	0.67	0.28*	0.76	0.45
6m	1m	118.7	106.7	37.1	63.2	1.21	0.43	0.47	0.33*	83.3	96.4	29.8	74.2	1.22	0.28	0.19*	0.22
	2-yr	86.7	67.4	26.9	36.8	1.18	0.68	0.59	0.38*	62.6	62.0	25.4	40.5	1.18	0.60	0.24*	0.27
	5y	74.2	50.1	24.3	24.6	1.12	0.91	0.64	0.48*	54.8	41.9	24.3	28.2	1.01	0.72	0.37	0.33*
	10y	65.4	41.3	21.5	16.9	1.00	0.78	0.89	0.65*	49.8	32.2	22.1	23.6	0.65	0.44*	0.66	0.44
12m	1m	118.7	106.7	54.7	63.4	1.24	0.73	0.53	0.42*	83.3	96.4	57.1	67.0	1.27	0.67	0.34*	0.38
	2m	86.7	67.4	45.0	40.3	1.24	1.03	0.61	0.47*	62.6	62.0	45.6	39.8	1.29	1.07	0.39*	0.40
	5m	74.2	50.1	44.5	28.4	1.08	1.17	0.60	0.57*	54.8	41.9	48.5	29.8	0.97	1.16	0.49	0.44*
	10m	65.4	41.3	41.4	18.4	1.07	1.23	0.84	0.77*	49.8	32.2	44.7	22.3	0.81	0.93	0.69	0.56*

Table B.1: Bond Yield Volatility Forecast Errors

Statistics of yield volatility forecast errors between model-implied and true forecasts for different horizons h and yield maturities m. μ and σ refer to the median values across simulated samples of the mean and standard deviations of the rolling-window forecasts. T refers to the RMedSE of the rolling-window forecasts. G_{na} and B_{na} refer to the NA Gaussian and Black models, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively. For the G_{na} model, we report the RMedSE between true and model-implied forecasts. For other models we report the RMedSE but scaled by the values reported for the G_{na} model in the first column. Forecast horizons are 3, 6, and 12 months and bond maturities are 6 months and 2, 5, and 9 years. The symbol * indicates the best performance for a given combination of horizon, maturity, and sample. The lower-bound sub-samples are collections of dates for which the 1-month rate is 25 bps or less.

		Entire Sample								ZLB Sample							
h	m	μ	σ	T	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	μ	σ	T	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
3-m	1-mth	375.1	312.2	87.3	10.8	0.93	0.93	0.93	0.75*	95.9	119.4	41.4	7.8	0.99	0.69	0.58*	0.69
	2-yr	368.8	278.5	73.3	7.9	0.97	0.96	0.92	0.81*	90.5	82.5	31.7	7.5	0.92	0.76	0.49*	0.61
	5-yr	382.1	251.9	63.1	7.7	1.04	1.01	0.91	0.78*	120.7	62.5	30.6	7.2	1.06	1.01	0.74	0.62*
	10-yr	417.9	217.1	53.4	7.9	1.01	0.92	0.93	0.73*	195.1	53.4	27.6	6.8	1.09	0.91	0.91	0.65*
6-m	1-mth	375.1	312.2	87.9	18.1	0.95	0.91	0.90	0.79*	95.9	119.4	37.3	14.2	0.96	0.68	0.53*	0.67
	2-yr	368.8	278.5	75.0	14.5	0.97	0.98	0.91	0.81*	90.5	82.5	27.8	13.1	0.94	0.87	0.58*	0.61
	5-yr	382.1	251.9	64.2	14.5	1.02	1.01	0.92	0.75*	120.7	62.5	26.9	13.3	1.05	1.02	0.79	0.58*
	10-yr	417.9	217.1	53.7	14.8	1.01	0.97	0.93	0.74*	195.1	53.4	27.8	12.9	1.07	0.94	0.94	0.68*
12-mth	1-mth	375.1	312.2	98.8	26.0	0.97	0.96	0.94	0.88*	95.9	119.4	32.4	24.0	0.90	0.67	0.47*	0.74
	2-yr	368.8	278.5	80.8	24.8	1.07	0.98	0.94	0.80*	90.5	82.5	32.1	23.4	0.83	0.87	0.66	0.54*
	5-yr	382.1	251.9	65.0	27.7	0.97	1.01	0.89	0.70*	120.7	62.5	37.3	24.1	0.99	0.97	0.81	0.61*
	10-yr	417.9	217.1	56.5	26.7	1.00	0.99	0.96	0.71*	195.1	53.4	39.3	21.3	1.05	1.00	1.04	0.70*

Table B.2: Bond Yield Forecast Errors

Statistics of yield forecast error between model-implied and true forecasts for different horizons h and yield maturities m. μ and σ refer to the median values across simulated samples of the mean and standard deviations of the rolling-window forecasts. T refers to the RMedSE of the rolling-window forecasts. G_{na} and B_{na} refer to the NA Gaussian and Black models, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively. For the G_{na} model, we report the RMedSE between true and model-implied forecasts. For other models, we report the RMedSE but scaled by the values reported for the G_{na} model in the first column. Forecast horizons are 3, 6, and 12 months and bond maturities are 6 months and 2, 5, and 9 years. The symbol * indicates the best performance for a given combination of horizon, maturity, and sample. The lower-bound sub-samples are collections of dates for which the 1-month rate is 25 bps or less.

It would seem that the same logic should apply to the B_{na} model. Turning to a direct comparison between the NA and ND Black style models, it is interesting to note that the B_{nd} model is the winner in lower-bound results for almost all horizon and maturity combinations. This difference in performance likely reflects the type of tension inherent in the B_{na} model (the tension between the scaling parameter S_n and the volatility parameters of the model). We note that the improvements of the B_{nd} model, relative to the B_{na} model, are more sizeable and concentrated in results close to the lower bound.

Biases and Correlations in Forecast Errors Table B.3 reports the bias in yield forecasts in Panel (a) and the bias in volatility forecasts in Panel (b). The first observation from Equation (B.3) is that the NA models have the largest full-sample biases, especially for longer maturities and longer horizons. By contrast, the SV-B_{nd} model has the smallest biases across full-sample results, except for three cases in Panel (a) and one case in Panel (b). The B_{nd} model offers mixed results. The biases in its volatility forecasts are closer to the flexible SV-B_{nd} model but the biases in its return forecasts are closer to the NA models.

Consider the G_{na} model. Its return forecast biases are large and negative but its volatility forecast biases are large and positive. From Equation (10), this may explain why the G_{na} Sharpe ratio estimates seem to offer accuracy close to the ND models while, at the same time, offering volatility forecasts that are much less accurate. In this case, the pattern in the biases compensates for the poor accuracy in the denominator of the Sharpe ratio.

Consider next the B_{na} model. Its return and volatility forecast biases are both large and negative. From Equation (10), this may explain why the B_{na} model's conditional Sharpe ratio forecasts offer the lowest accuracy while, at the same time, offering yield forecasts and volatility forecasts that are more accurate than the G_{na} model, and not so different from that of the B_{nd} model. In this case, the pattern in the biases substantially aggravates the accuracy of Sharpe ratio estimates. Finally, Equation (10) suggests that the correlation between yield forecasts and volatility forecasts can also affect the accuracy of Sharpe ratio estimates. In unreported results, we find that the correlations are generally negative but relatively small (e.g., between around -0.2 and 0.2 for NA models).

VI Equation 10

Define the forecast error $fe_j(x)$ and the mean squared forecast error $msfe_j(x)$ for some forecast x from model j:

$$\begin{aligned} fe_j(x) &\equiv (x_j - x_0) \\ msfe_j(x) &\equiv E_0[fe_j(x)^2] = E[(x_j - x_0)^2 | x_0], \end{aligned}$$

where the expectation operator $E_0[\cdot]$ conditions on the known moments from the simulation, x_j is the forecast from model j and x_0 is the know true forecast. We ignore the time subscript for simplicity. We have that:

$$\begin{aligned} fe_j(SR) =& \frac{xr_j}{\sigma_j} - \frac{xr_0}{\sigma_0} = \frac{\sigma_0 xr_j - \sigma_j xr_0}{\sigma_j \sigma_o} \\ =& \frac{1}{\sigma_j} \left(fe_j(xr) - SR_0 fe_j(\sigma) \right), \end{aligned} \tag{B.27}$$

so that

$$fe_j(SR)^2 = \frac{1}{\sigma_j^2} \left(fe_j(xr)^2 + SR_0^2 fe_j(\sigma)^2 - 2SR_0^2 fe_j(xr) fe_j(\sigma) \right). \tag{B.28}$$

(a) Return Forecast Bias

h	m	Entire Sample					ZLB Sample				
		G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
3-mth	1-mth	-1.1	-1.1	-1.6	0.6*	3.2	-2.2	−0.5*	-3.2	3.4	3.7
	2-yr	-3.5	-4.0	-3.9	-2.2	0.6*	-4.3	-4.3	-4.1	0.1*	1.2
	5-yr	-5.3	-5.4	-6.3	-4.5	−1.6*	-4.8	-5.4	-5.8	-2.0	0.1*
	10-yr	-6.0	-5.3	-6.1	-5.3	−2.7*	-5.2	-5.5	-4.9	-3.2	−1.5*
6-m	1-mth	-2.4	-2.5	-3.3	0.2*	6.2	-3.2	−0.6*	-4.5	4.5	7.7
	2-yr	-7.3	-7.3	-8.3	-5.1	0.1*	-7.5	-7.0	-8.5	−0.3*	1.9
	5-yr	-10.8	-10.4	-12.1	-8.8	−4.1*	-8.8	-9.5	-10.2	-3.7	−0.7*
	10-yr	-12.0	-10.4	-11.8	-10.4	−6.3*	-9.4	-10.3	-9.6	-7.3	−3.6*
12-mth	1-mth	-8.4	-8.0	-8.5	−2.9*	6.7	-9.3	−4.0*	-8.4	5.7	13.1
	2-yr	-16.4	-16.2	-17.2	-11.9	−2.7*	-12.9	-12.8	-16.1	−1.7*	3.0
	5-yr	-22.9	-20.3	-23.7	-18.8	−7.8*	-17.0	-17.3	-19.8	-10.0	−2.5*
	10-yr	-22.2	-20.5	-22.3	-21.1	−12.0*	-17.3	-18.3	-19.6	-14.7	−7.6*

(b) Volatility Forecast Bias

h	m	Entire Sample					ZLB Sample				
		G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}	G_{na}	$A1_{na}$	B_{na}	B_{nd}	SV-B_{nd}
3-m	1-mth	37.3	47.0	−12.0*	-2.4	-0.3	68.9	82.0	−6.5*	-0.5	-1.7
	2-yr	21.1	26.5	−14.4*	2.6	1.6	35.8	42.0	−12.7*	0.8	0.3
	5-yr	12.8	15.3	−13.9*	6.0	3.0	24.2	23.5	−10.5*	8.2	2.5
	10-yr	9.0	8.7	−5.6*	7.9	3.8	15.3	10.2	−2.6*	10.3	3.5
6-m	1-mth	40.5	54.7	−23.9*	-7.9	-4.0	74.2	90.9	−19.7*	-7.4	-7.0
	2-yr	25.2	32.4	−25.1*	0.0	0.6	40.4	47.7	−24.4*	-3.4	-3.3
	5-yr	14.0	19.9	−20.9*	4.6	2.5	25.2	25.9	−20.3*	4.0	-0.2
	10-yr	10.7	9.3	−9.8*	8.0	3.8	16.6	9.2	−8.7*	8.6	1.4
12-mth	1-mth	37.6	56.5	−44.4*	-14.7	-8.6	67.0	85.1	−44.8*	-21.2	-16.2
	2-yr	23.8	34.6	−40.2*	-6.3	-1.2	38.6	49.8	−42.5*	-12.4	-9.5
	5-yr	14.6	21.2	−32.5*	1.3	0.8	25.7	26.4	−34.7*	-3.7	-5.2
	10-yr	9.1	10.1	−18.8*	5.2	1.2	13.9	6.1	−19.7*	3.7	-3.6

Table B.3: Return Forecast Bias

Median bias (bps) between model-implied forecasts and true forecasts for returns in Panel (a) and volatility in Panel (b). G_{na} and B_{na} refer to the NA Gaussian and Black models, respectively. B_{nd} and SV-B_{nd} refer to the ND Black models with and without BEKK volatility, respectively. Forecast horizons are 3, 6, and 12 months and maturities are 1 month and 2 years, 5 years, and 10 years. The symbol * indicates the best performance for each combination of horizon, maturity, and sample. The lower-bound sub-samples are defined as the collections of dates for which the 1-month rate is 25 bps or less.

Suppose that time-t squared Sharpe ratio forecast errors $fe_j(SR)^2$ are conditionally independent from time-t variance of yields σ_j^2. Then, we have that:

$$msfe_j(SR) = \frac{1}{E_0[\sigma_j^2]} \left(msfe_j(xr) + SR_0^2 msfe_j(\sigma) - 2SR_0^2 E_0[fe_j(xr) fe_j(\sigma)] \right)$$

$$= \frac{1}{E_0[\sigma_j^2]} \left(msfe_j(xr) + SR_0^2 msfe_j(\sigma) + 2SR_0 E_0[fe_j(xr)] E_0[fe_j(\sigma)] - 2SR_0 cov(fe_j(xr), fe_j(\sigma)) \right), \quad \text{(B.29)}$$

which yields to Equation (10) with the appropriate coefficient definitions. Therefore, biases that have the same signs or forecast errors that are negatively correlated increase the Sharpe ratio forecast errors for model j.