Term Structure Analysis with Big Data: One-Step Estimation Using Bond Prices

Martin M. Andreasen[†]

Jens H. E. Christensen[‡]

Glenn D. Rudebusch[*]

Abstract

Nearly all studies that analyze the term structure of interest rates take a *two-step* approach. First, actual bond prices are summarized by interpolated synthetic zero-coupon yields, and second, some of these yields are used as the source data for further empirical examination. In contrast, we consider the advantages of a *one-step* approach that directly analyzes the universe of bond prices. To illustrate the feasibility and desirability of the one-step approach, we compare arbitrage-free dynamic term structure models estimated using both approaches. We also provide a simulation study showing that a one-step approach can extract the information in large panels of bond prices and avoid any arbitrary noise introduced from a first-stage interpolation of yields.

JEL Classification: C55, C58, G12, G17

Keywords: extended Kalman filter, fixed-coupon bond prices, arbitrage-free Nelson-Siegel model

---

An earlier draft of this paper was circulated under the title "Term Structure Analysis with Big Data." We thank the editor and two anonymous referees for helpful comments and suggestions. We also thank participants at the Big Data in Dynamic Predictive Econometric Modeling Conference and the 10th Annual SoFiE Conference for helpful comments. Finally, we thank Patrick Shultz for outstanding research assistance. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

[†]Aarhus University, CREATES, and the Danish Finance Institute. Address: Fuglesangs Allé 4, Aarhus V, 8210, Denmark; e-mail: mandreasen@econ.au.dk.

[‡]Corresponding author: Federal Reserve Bank of San Francisco, 101 Market Street MS 1130, San Francisco, CA 94105, USA; phone: 1-415-974-3115; e-mail: jens.christensen@sf.frb.org.

[*]Federal Reserve Bank of San Francisco, www.glennrudebusch.com; e-mail: glenn.rudebusch@sf.frb.org.

This version: April 19, 2018.

1 Introduction

Most term structure analysis takes a two-step approach when investigating the pricing of fixed-income securities—although the initial step is usually outsourced and taken for granted. In that first step, constant-maturity zero-coupon yields are constructed from the universe of coupon bond prices. In the second step, some of these synthetic yields are used as an input to estimate a dynamic term structure model (DTSM) (see Dai and Singleton (2000), Duffee (2002) among many others). This two-step process arose decades ago because of the computational burden of working directly with large data sets of actual bond prices. However, the convenience of this two-step approach has also resulted in relatively little interest or attention being given to the construction of the synthetic zero-coupon yields. This situation has continued despite the documented challenges and problems in constructing synthetic zero-coupon yields (see, for instance, Bliss (1997) and Gürkaynak et al. (2007, 2010)). Even more concerning, some researchers such as Dai et al. (2004) and Fontaine and Garcia (2012) have argued that using synthetic interpolated yields can erase interesting bond price dynamics by excessively smoothing the data and may even introduce unnecessary measurement errors.

The contribution of the present paper is to show that the initial step of constructing synthetic zero-coupon yields can be avoided, as progress in computing power now allows researchers to work directly with the bigger data universe of coupon bonds. We illustrate this alternative by comparing identical DTSMs that are estimated by the one- and two-step approaches—both on empirical and simulated samples of coupon bond prices.

Our empirical case study for comparing the one- and two-step estimation approaches is the Canadian government bond market between January 2000 and April 2016. The Canadian market is a good laboratory for our analysis for several reasons. First, Canadian bonds face no appreciable credit risk during our sample and do not attract the same large and variable liquidity and safety premiums that affect the pricing of U.S. Treasuries.[1] Second, Canadian bond yields spent relatively little time near the zero lower bound—again compared to the United States—which simplifies our analysis. Finally, the number of Canadian government bonds is representative of sovereign bond markets in many developed countries. In total, our Canadian sample for the one-step DTSM estimation contains end-of-month prices on 105 bonds. The corresponding data for the two-step approach follows the existing literature and uses a limited number of synthetic zero-coupon yields. We consider two sources for these synthetic yields. The first data set is produced by the Bank of Canada and described in

[1] For example, in constructing their interpolated nominal U.S. Treasury yield curves, Gürkaynak et al. (2007) generally exclude the two most recently issued securities, i.e. the "on-the-run" and "first off-the-run" bonds, which often trade at a premium. A one-step approach could also exclude these bond prices or augment the DTSM of interest to accommodate bond-specific liquidity characteristics as in Fontaine and Garcia (2012) and Andreasen et al. (2018), but it would complicate our analysis.

Bolder et al. (2004).[2] We construct the second data set of synthetic yields by estimating the flexible parametric discount function of Svensson (1995) month-by-month on the same panel of coupon bonds as used for the one-step approach.[3] The differences between these two data sets of synthetic zero-coupon yields are generally small for maturities within the one- to twenty-year maturity range, but the differences may easily exceed ten basis points outside this maturity range where fewer bonds are available. This observation provides tentative evidence that the various curve-fitting techniques may induce non-negligible measurement errors in synthetic yields.

We then estimate the same DTSM on Canadian bond prices via the one- and two-step approaches. Our benchmark DTSM is the arbitrage-free Nelson-Siegel (AFNS) model of Christensen et al. (2011), which is a Gaussian affine model where a level, slope, and curvature factor explain the evolution of the yield curve. We find that the one-step approach gives a substantially closer fit to the underlying coupon bond data than the conventional two-step approach. For the AFNS model, the fit to market prices of coupon bonds may deteriorate by as much as 41% when going from a one-step to a two-step approach, which may add considerable noise to predicted bond prices from an estimated DTSM. We also find that the parameters determining the functional form between bond yields and the factors (i.e., the risk-neutral parameters) are those that are mostly affected by the choice of estimation approach.

To complement these empirical estimates, we also explore the finite-sample properties of the one- and two-step approaches in a Monte Carlo study. A novel feature of this simulation experiment is to work at the level of coupon bonds and hence account for estimation uncertainty in the construction of synthetic zero-coupon yields within the two-step approach. The main insight from this Monte Carlo study is that DTSMs may be estimated more reliably by using observed bond prices instead of synthetic zero-coupon yields. Although these synthetic yields are estimated very accurately with well-established curve-fitting techniques, we nevertheless find that seemingly negligible errors in these synthetic yields do affect the estimated parameters in a DTSM. In particular, all risk-neutral parameters are estimated with smaller biases and greater precision in the one-step approach compared to the conventional two-step approach.

Finally, to showcase the suitability, tractability, and general applicability of the one-step approach, we perform an out-of-sample forecast exercise of the three-month and ten-year Canadian bond yield. In addition to the AFNS model, we also include a nonlinear DTSM that

[2]See Diez de los Rios (2015) for an empirical application using these data.

[3]This is the exact procedure used by Gürkaynak et al. (2007) and many others. Alternative functional forms could be considered such as the cubic splines used by Steeley (2008), the hybrid combination of cubic splines and parametric functions advocated by Faria and Almaida (2017), or the optimally smooth spline yield curves derived from an exact bootstrap method based on the Moore-Penrose pseudoinverse developed by Filipović and Willems (2016).

enforces the zero lower bound and a five-factor model to get an even tighter fit of long-term Canadian bonds than implied by a three-factor model. We find that the one-step approach has either the best or nearly-the-best forecast accuracy for the three models considered. The performance of the two-step approach can vary depending on which synthetic zero-coupon yields are used for estimating the models.

This paper is most closely related to the work of Fontaine and Garcia (2012) and Pancost (2018), which are among the few papers that also use the underlying data on coupon bond prices to estimate DTSMs. Fontaine and Garcia (2012) consider pairs of old and newly issued bonds within various maturity buckets to study the on-the-run liquidity premium in U.S. Treasuries. Pancost (2018) uses the full sample of U.S. Treasuries to show that the pricing errors on bonds are predictable over time and in the cross section, and that realized excess bond returns can be used to improve the estimation of the time series parameters in a DTSM. Importantly, neither Fontaine and Garcia (2012) nor Pancost (2018) compare their results to those obtained from a corresponding two-step approach, which is a key contribution of the present paper.

The remainder of the paper is structured as follows. Section 2 describes the Canadian government bond data, while Section 3 summarizes the AFNS model and presents its estimation results on Canadian data. Section 4 is devoted to our Monte Carlo study, while Section 5 provides an out-of-sample forecasting exercise of Canadian bond yields. Section 6 concludes. Appendices available online contain additional details related to the paper.

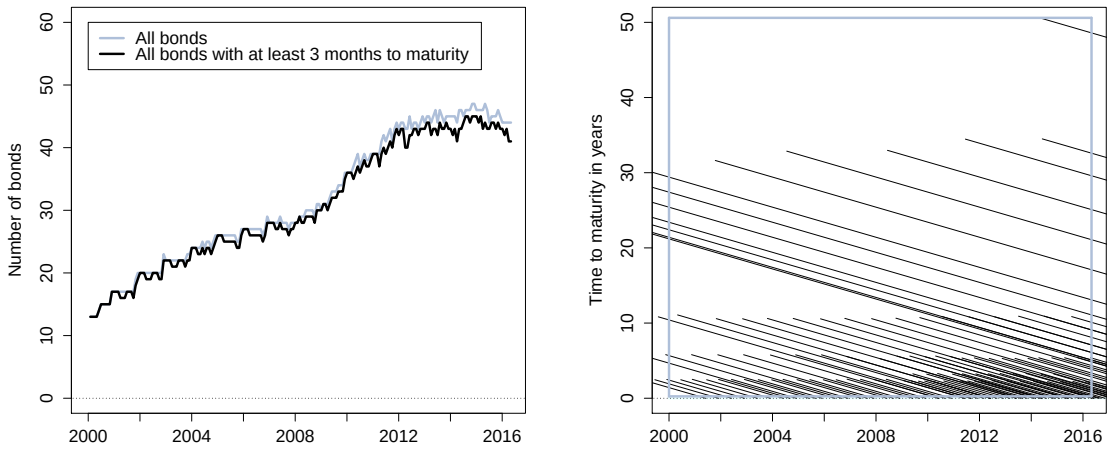
2 The Canadian Bond Market

This section describes the market for Canadian government bonds. We first describe the universe of Canadian bonds, which will be used for the one-step approach. Then we present two data sets of synthetic zero-coupon yields for the two-step approach.

2.1 The Universe of Government Bonds

As of April 2016, the Canadian government bond market had a total outstanding notional amount of CAD 512.5 billion, which is equivalent to 25% of the gross domestic product in Canada. The Canadian government holds a AAA rating with a stable outlook by all major rating agencies, meaning that no correction for credit risk is required. Panel (a) of Figure 1 shows that the number of coupon bonds grows gradually from about 15 bonds at the start of the sample to roughly 45 bonds in 2012, where it has remained until the end of our sample in 2016. The size of the Canadian market is representative of sovereign bond markets in several developed countries.[4]

[4]Online Appendix A contains the corresponding details for France, Switzerland, and the U.K.



(a) Number of Canadian bonds

(b) Maturity distribution of Canadian bonds

Figure 1: **Description of the Canadian Bond Market**
Panel (a) shows the number of Canadian government bonds at each date. The solid grey line refers to the entire sample of bonds. The solid black line indicates the number of securities after eliminating bonds with less than three months to maturity. Panel (b) shows the maturity distribution of the full set of Canadian government bonds in our sample. The grey rectangle indicates the subsample used throughout the paper.

The time-varying maturity distribution of all 105 bonds in our sample is illustrated in panel (b) of Figure 1, where each security is represented by a downward-sloping line showing its remaining years to maturity at each date. The short end of the bond market has been densely populated with many two-year bonds. Five- and ten-year bonds were issued fairly regularly since the start of our sample. At the very long maturities, thirty-year bonds were issued approximately every three years, and a single fifty-year bond was issued in 2014.[5] All bond prices are represented by their clean mid-market price as provided by Bloomberg. Following Gürkaynak et al. (2007), securities with less than three months to maturity are excluded from our sample, as the implied yield on these securities often display erratic behavior.

2.2 Synthetic Zero-Coupon Yields

The data for the two-step approach follows the existing literature, which represents the universe of bonds by a limited number of synthetic zero-coupon yields. We consider two sources for such synthetic yields. The first data set is produced by the Bank of Canada using the "Merrill Lynch exponential spline model" and is publicly available.[6] We construct the second

[5]The contractual characteristics of all 105 bonds and the number of monthly observations for each bond are reported in online Appendix A.

[6]See Bolder et al. (2004) for a description of the yield curve construction and the algorithm used to filter out "strange" observations.

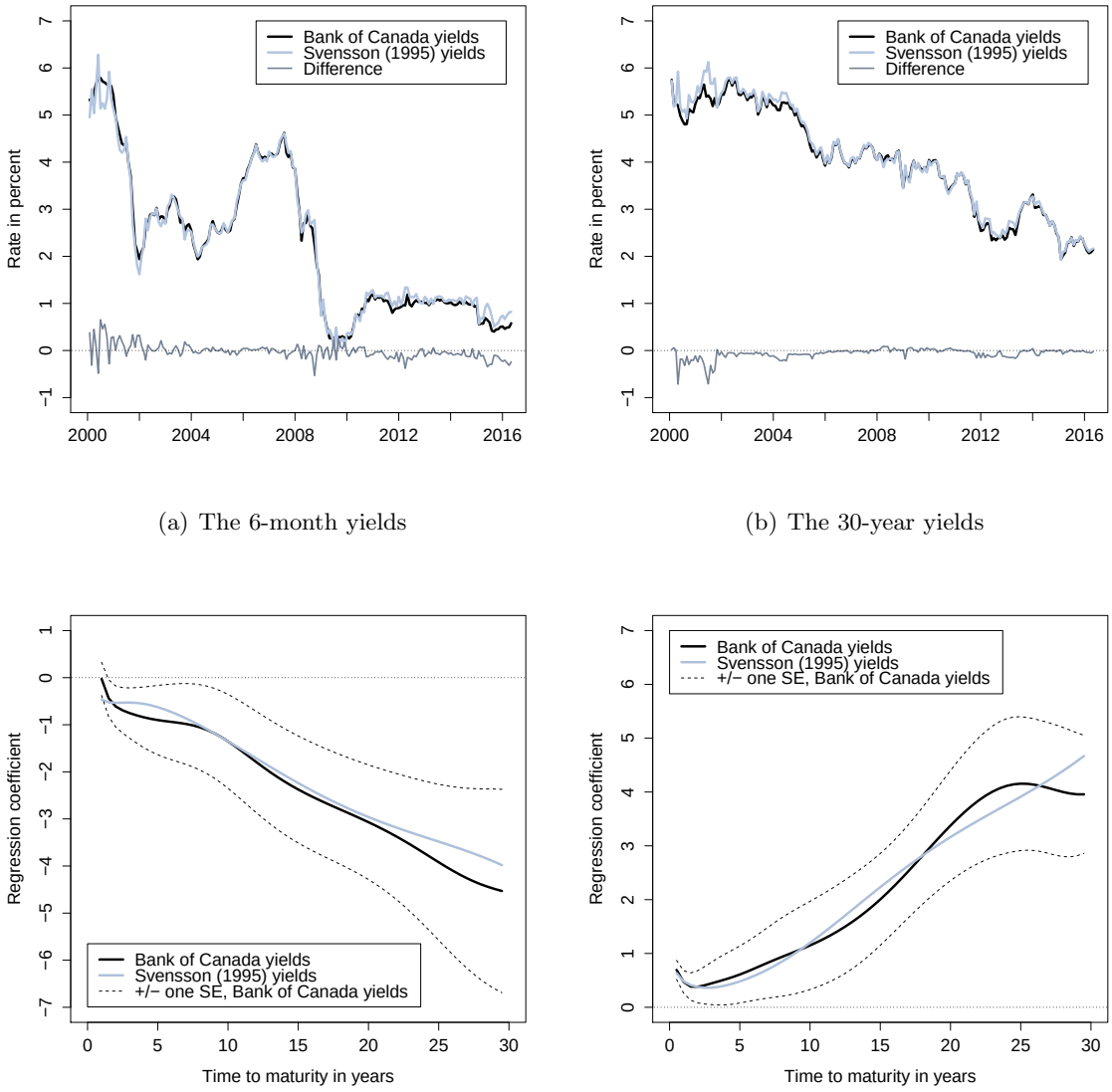
Maturity in months	Mean diff.	Mean abs. diff.	Max. abs. diff.	Correlation	
				Levels	Diff.
3	0.78	21.52	105.24	0.982	0.410
6	-1.95	11.41	65.25	0.995	0.693
12	-3.80	4.77	22.13	0.999	0.966
24	-1.13	3.22	15.85	1.000	0.986
36	1.12	2.69	11.74	1.000	0.990
60	1.42	3.25	23.37	1.000	0.992
84	-0.71	4.85	21.57	0.999	0.989
120	-5.37	5.48	19.46	1.000	0.988
240	5.12	5.84	20.03	0.999	0.968
360	-6.63	7.86	71.43	0.995	0.848

Table 1: **Comparing Two Data Sets of Synthetic Zero-Coupon Yields**
The table reports the summary statistics for the mean differences, the mean absolute differences, and the maximum absolute differences between synthetic Canadian zero-coupon yields from the Bank of Canada and our implementation of the Svensson (1995) discount function. These differences are reported in annual basis points. The last two columns report the correlations between the two yield series for each maturity in levels and first differences, respectively. The data series are monthly covering the period from January 31, 2000, to April 30, 2016.

data set by estimating the flexible discount function of Svensson (1995) month-by-month on the same panel of coupon bonds as used for the one-step approach (see online Appendix B). This technique is quite common worldwide as a procedure to create zero-coupon yields. For each data set, we extract synthetic yields with ten maturities of 0.25, 0.5, 1, 2, 3, 5, 7, 10, 20, and 30 years.

Table 1 reports summary statistics for the differences between the two data sets at various maturities. The mean absolute differences for yields in the one- to twenty-year maturity range are within six basis points and hence small. Larger deviations emerge at the very short and long maturities with mean absolute differences at the six-month and thirty-year maturities of 11 and 8 basis points, respectively. Large maximum outlier differences are also evident. Non-negligible discrepancies are also evident in the correlations between the two data sets shown in the last two columns in Table 1. The correlations are clearly less than one at short and long maturities. To further illustrate these differences, Figure 2 plots the six-month and thirty-year yields from the two data sets. There are notable differences at the start of the sample and when the short rate approaches the zero lower bound in 2009.

Another way to evaluate the magnitude of these differences is to re-visit two classic regressions. The first is due to Campbell and Shiller (1991), where realized returns are regressed on the slope of the yield curve. Panel (c) in Figure 2 shows that the loadings in these regressions differ quite a bit at the short and long end of the yield curve but are almost identical in the five- to twenty-year maturity spectrum. The second regression is due to Fama (1976),



(a) The 6-month yields

(b) The 30-year yields

(c) Loadings in Campbell-Shiller regressions

(d) Loadings in forward rate regressions

Figure 2: **Two Data Sets of Synthetic Zero-Coupon Yields: Key Differences**
Panel (a) shows the six-month synthetic yields from the Bank of Canada and our implementation of the Svensson (1995) discount function. Panel (b) shows the thirty-year synthetic yields from the Bank of Canada and our implementation of the Svensson (1995) discount function. Panel (c) shows δ_k from the regression $y_{t+h}(k-h)-y_t(k) = \alpha_k+\delta_k \frac{h}{k-h}\left(y_t(k) - y_t(h)\right)+\varepsilon_t(k)$ with $h = 6$ months, where $y_t(k)$ refers to the yield in period t with k months to maturity. Panel (d) shows $\theta(k)$ in the regression $xhpr_{t+h}(k) = \mu(k) + \theta_k x_t(k) + \nu_{t+h}(k)$ with $h = 6$ months, where $xhpr_{t+h}(k) \equiv hpr_{t+h}(k) - \frac{h}{12}y_t(h)$ is the excess holding period return and $hpr_{t+h}(k) \equiv -\frac{k-h}{12}y_{t+h}(k-h) + \frac{k}{12}y_t(k)$ is the holding period return. The variable $x_t(k)$ denotes the forward spread $f_t^{(k-h,k)} - \frac{h}{12}y_t(h)$, where $f_t^{(k-h,k)} \equiv \frac{k}{12}y_t(k) - \frac{k-h}{12}y_t(k-h)$ is the forward rate between time $t+k-h$ and $t+k$.

where realized excess returns are regressed on the slope of the forward curve. Although most regression loadings in Panel (d) coincide closely, we do find substantial differences beyond the

twenty-year maturity. Here, loadings increase monotonically for the Svensson (1995) yields but not when using the synthetic yields from the Bank of Canada. Importantly, though, these differences are not statistically significant across the two regressions, as the estimated regression loadings based on the Svensson (1995) yields are well within one standard deviation of the estimated coefficients from the Bank of Canada yields.

3 Empirical Application

This section presents an empirical application of the one- and two-step approaches. We first present a benchmark DTSM, and then describe the econometric issues related to the one- and two-step estimation approaches. After discussion of the results, we consider several robustness checks as well.

3.1 A Gaussian DTSM

To model the bond market, we use the three-factor Gaussian DTSM of Christensen et al. (2011), which can be viewed as a restricted version of the affine DTSMs in Dai and Singleton (2000). In this arbitrage-free Nelson-Siegel (AFNS) model, the state vector is denoted by $X_t = (L_t, S_t, C_t)$, where L_t, S_t, and C_t are the level, slope and curvature factors. The instantaneous risk-free rate is defined as $r_t = L_t + S_t$, and the risk-neutral (or $\mathbb{Q}$-) dynamics of the state variables are given by

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\lambda & \lambda \\ 0 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} dt + \Sigma \begin{pmatrix} dW_t^{L,\mathbb{Q}} \\ dW_t^{S,\mathbb{Q}} \\ dW_t^{C,\mathbb{Q}} \end{pmatrix}. \tag{1}$$

Here, $dW^{i,\mathbb{Q}}$ for $i = \{L, S, C\}$ denotes independent Wiener processes and Σ is a constant covariance matrix with dimensions 3×3.[7] The zero-coupon bond yield at maturity τ is

$$y(\tau; X_t) = L_t + \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau}\right) S_t + \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) C_t - \frac{A(\tau)}{\tau}, \tag{2}$$

where $A(\tau)$ is a convexity term that adjusts the functional form in Nelson and Siegel (1987) to ensure absence of arbitrage (see Christensen et al. (2011))

The model is closed by adopting the essentially affine specification for the market price of risk Γ_t of Duffee (2002). That is, we let $\Gamma_t = \gamma^0 + \gamma^1 X_t$, where $\gamma^0 \in \mathbf{R}^3$ and $\gamma^1 \in \mathbf{R}^{3\times 3}$ contain unrestricted parameters. The physical (or $\mathbb{P}$-) dynamics of the three factors in the

[7] As discussed in Christensen et al. (2011), the unit root in the level factor implies that the model is only free of arbitrage for bonds with a finite horizon. For our sample of Canadian bonds described in Section 2, and most other sovereign bond markets, this restriction is not binding and therefore of no practical relevance.

AFNS model are therefore

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & \kappa_{12}^{\mathbb{P}} & \kappa_{13}^{\mathbb{P}} \\ \kappa_{21}^{\mathbb{P}} & \kappa_{22}^{\mathbb{P}} & \kappa_{23}^{\mathbb{P}} \\ \kappa_{31}^{\mathbb{P}} & \kappa_{32}^{\mathbb{P}} & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left(\begin{pmatrix} \theta_1^{\mathbb{P}} \\ \theta_2^{\mathbb{P}} \\ \theta_3^{\mathbb{P}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right) dt + \Sigma \begin{pmatrix} dW_t^{L,\mathbb{P}} \\ dW_t^{S,\mathbb{P}} \\ dW_t^{C,\mathbb{P}} \end{pmatrix}, \tag{3}$$

where $\kappa_{i,j}^{\mathbb{P}}$ and $\theta_i^{\mathbb{P}}$ are free parameters, subject to X_t being stationary under the $\mathbb{P}$-measure.

3.2 Estimation Methodology for One- and Two-Step Approaches

To describe the econometric implementation of the one-step approach, let $P_t^i(\tau, C)$ denote the price at time t of the ith coupon bond, which matures at time $t+\tau$ and pays the coupon C semi-annually at times t_j. In the absence of arbitrage, the clean price of this coupon bond must equal the discounted sum of all remaining payments, i.e.,

$$P_t^i(\tau, C) = \frac{C}{2}\frac{(t_1 - t)}{1/2} P_t^{\text{zc}}(t_1 - t) + \sum_{j=2}^{N} \frac{C}{2} P_t^{\text{zc}}(t_j - t) + P_t^{\text{zc}}(t_N - t), \tag{4}$$

where $t < t_1 < \ldots < t_N = \tau$. Here, $P_t^{\text{zc}}(\tau) = \exp\{-y(\tau; X_t)\tau\}$ denotes the price of the zero-coupon bond with τ years to maturity, and $y(\tau; X_t)$ is the zero-coupon yield from the DTSM.[8] The corresponding bond price in the data is denoted $P_t^{i,Data}(\tau, C)$. A preliminary analysis showed that the pricing errors ε_t^i of the AFNS model tend to be larger for longer-term bonds. Therefore, we scale these errors by duration, so the measurement equation for the ith bond price in the one-step approach is given by

$$P_t^{i,Data}(\tau, C) = P_t^i(\tau, C) + D_t^{i,Data}(\tau, C)\varepsilon_t^i, \tag{5}$$

where we apply the model-free Macaulay duration $D_t^{i,Data}(\tau, C)$.[9] The pricing errors are assumed to be Gaussian, independent across time, and independent to the state innovations in equation (3), i.e. $\varepsilon_t^i \sim \mathcal{NID}\left(0, \sigma_\varepsilon^2\right)$. The state transition dynamics for X_t under the $\mathbb{P}$-measure is given by equation (3).

The states X_t are taken to be unobserved and must be estimated along with the model

[8] The continuous-time formulation of our model makes the implementation of equation (4) straightforward. For discrete-time models with one period exceeding one day (say, a week or a month), standard interpolation schemes may be used to price the coupon payments related to the ith bond at time t. Note that time in equation (4) is measured in years, meaning that $t_1 - t$ must be divided by 0.5 to obtain the fraction of the semi-annual coupon payment $C/2$ that remains between time t and t_1.

[9] A similar scaling is used in Gürkaynak et al. (2007) and Hu et al. (2013) to estimate the static discount function of Svensson (1995). Equation (5) is equivalent to dividing bond prices by duration to obtain a first-order approximation of the implied yield to maturity on a coupon bond. Using the exact yield to maturity in the estimation is computationally much more demanding because it requires solving a nonlinear fixed-point problem for every evaluation of the measurement equation, although we do discuss the results for this alternative specification in Section 3.4.

parameters ψ from the panel of bond prices. The nonlinear relationship between X_t and the price of a coupon bond $P_t^i(\tau, C)$ in equation (4) implies that our arbitrage-free DTSM cannot be estimated using the Kalman filter. Instead, the extended Kalman filter (EKF) is used to obtain an approximated log-likelihood function $L^{EKF}(\psi)$, which serves as the basis for estimating ψ by quasi-maximum likelihood (QML), as described in further detail in the online Appendix C.

The econometric implementation of the two-step approach is well-known but summarized here for completeness. Let the synthetic zero-coupon yields in the data be denoted by $y_t^{Data}(\tau)$, and let $y(\tau, X_t)$ denote the corresponding yield from the DTSM. The measurement equation is

$$y_t^{Data}(\tau) = y(\tau, X_t) + \varepsilon_t(\tau),$$

for a selection of constant maturities yields, as indexed by τ. Here, $\varepsilon_t(\tau) \sim \mathcal{NID}\left(0, \sigma_\varepsilon^2\right)$ and accounts for estimation errors in the construction of these synthetic yields and pricing errors in the DTSM. The state transition dynamics for X_t under the $\mathbb{P}$-measure is given by equation (3). For the AFNS model, the zero-coupon yields are affine in X_t as seen from equation (2), and all model parameters ψ are therefore estimated by maximum likelihood (ML) using the Kalman filter.

3.3 Estimation Results

The estimated model parameters in the AFNS model are reported in Table 2 when using the one- and two-step approaches. The conventional two-step approach is implemented on the two samples of synthetic yields discussed in Section 2.2 to explore whether the highlighted differences in the two data sets affect the estimated parameters. Hence, the one-step approach uses all available bond prices with maturities exceeding three months, whereas the two-step approach only uses the ten maturities selected in Table 1. In the interest of simplicity, we focus on the most parsimonious version of the AFNS model with independent factor dynamics. This restriction comes at practically no loss of generality for the reported results as the estimated factors and model fit are insensitive to omitting the off-diagonal terms in $\mathcal{K}^{\mathbb{P}}$ and Σ.[10]

We first note that the estimates of $\mathcal{K}^{\mathbb{P}}$ and $\theta^{\mathbb{P}}$ have sizeable standard errors in all three data sets, which is a well-known characteristic of estimating persistent autoregressive processes using a relatively short time span. The diagonal elements in Σ and λ are estimated much more accurately and reveal some notable differences. First, the volatility of the level factor σ_{11} is 0.0071 in the two-step approach based on yields from Bank of Canada, but only 0.0052 in the one-step approach and in the two-step approach based on Svensson (1995) yields. Second,

[10] See for instance Christensen et al. (2011), who also show that this restricted model often does better at forecasting yields out of sample than the most flexible version of the AFNS model, where $\mathcal{K}^{\mathbb{P}}$ and Σ are unrestricted.

Par.	One-step approach		Two-step approach			
			Bank of Canada yields		Svensson (1995) yields	
	Est	SE	Est	SE	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.1060	0.0763	0.2172	0.3086	0.0835	0.1327
$\kappa_{22}^{\mathbb{P}}$	0.2157	0.1443	0.1839	0.1696	0.2982	0.1969
$\kappa_{33}^{\mathbb{P}}$	0.7255	0.3649	0.4214	0.2675	0.3543	0.2301
σ_{11}	0.0052	0.0001	0.0071	0.0001	0.0052	0.0001
σ_{22}	0.0103	0.0010	0.0085	0.0005	0.0103	0.0004
σ_{33}	0.0207	0.0015	0.0197	0.0013	0.0212	0.0013
$\theta_1^{\mathbb{P}}$	0.0529	0.0034	0.0542	0.0111	0.0477	0.0143
$\theta_2^{\mathbb{P}}$	-0.0275	0.0093	-0.0295	0.0136	-0.0251	0.0088
$\theta_3^{\mathbb{P}}$	-0.0230	0.0060	-0.0187	0.0129	-0.0181	0.0156
λ	0.3747	0.0105	0.3070	0.0047	0.4511	0.0051

Table 2: **Parameter Estimates in the AFNS Model**
This table reports the estimated parameters (Est) in the AFNS model with independent factors and their standard errors (SE) using either the one-step or two-step approach. The SE in the one-step approach are computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, which we compute as outlined in Harvey (1989). The SE in the two-step approach are computed from the inverse of the variance of the score. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

the volatility of the slope factor σ_{22} is 0.0085 in the two-step approach using yields from the Bank of Canada, whereas we find $\sigma_{22} = 0.0103$ in the two other data sets. Finally, the decay parameter λ is 0.375 in the one-step approach, 0.305 in the two-step approach based on Bank of Canada yields, and 0.451 in the two-step approach using Svensson (1995) yields. These findings reveal that the estimated parameters in a DTSM are more affected by the choice of synthetic yields data set than by the use of bond prices. Clearly, even small differences between synthetic yields of the same maturity can alter the estimation results.

Table 3 evaluates the ability of the three estimated AFNS models to match coupon bond prices. The pricing errors are computed using the implied yield to maturity on each coupon bond to make the errors comparable across securities. That is, for the price on the ith coupon bond $P_t^i(\tau, C)$, we find the value of $y_t^i(\tau, C)$ that solves

$$P_t^i(\tau, C) = \frac{C}{2}\frac{(t_1 - t)}{1/2} e^{-y_t^i(\tau,C)(t_1-t)} + \sum_{j=2}^{N} \frac{C}{2} e^{-y_t^i(\tau,C)(t_j-t)} + e^{-y_t^i(\tau,C)\tau}. \tag{6}$$

For the model-implied estimate of this bond price $\hat{P}_t^i(\tau, C)$ we find the corresponding yield $\hat{y}_t^i(\tau, C)$ and report the pricing errors as $y_t^i(\tau, C) - \hat{y}_t^i(\tau, C)$. Table 3 shows that the two-step approach provides a fairly tight fit to the underlying coupon bond prices with an overall root mean squared error (RMSE) of 8.31 basis points for the Bank of Canada yields and 7.90 basis

Maturity bucket in years	No. obs.	DTSM: The AFNS model						Static model: The Svensson (1995) discount function	
		One-step approach		Two-step approach					
				Bank of Canada yields		Svensson (1995) yields			
		Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
0-2	1,472	-0.09	5.75	2.40	7.87	0.33	9.81	-1.08	8.83
2-4	1,098	0.46	4.76	1.78	7.20	2.02	5.61	0.87	4.27
4-6	744	-0.32	4.01	1.04	4.37	-1.23	4.61	0.40	3.50
6-8	404	-1.24	5.46	0.04	5.02	-3.12	6.61	-1.89	4.50
8-10	477	-2.54	6.07	-2.27	6.61	-4.62	7.81	-2.95	5.43
10-12	289	-1.07	6.36	-1.09	8.56	-2.01	8.35	-2.06	5.84
12-14	155	3.78	6.72	4.70	11.98	4.79	11.18	2.01	3.65
14-16	168	0.76	4.32	-1.61	9.05	0.36	8.28	0.35	2.87
16-18	179	0.70	4.66	-1.97	9.89	0.94	8.84	0.24	3.80
18-20	192	1.71	4.33	2.02	8.89	4.88	8.64	0.68	3.60
20-22	186	3.64	5.98	5.06	10.06	7.36	10.32	2.32	4.58
22-24	142	0.84	5.74	2.37	7.27	4.62	7.09	1.39	3.59
24-26	124	0.05	5.63	3.75	8.37	4.67	7.43	1.63	3.56
26-28	113	-5.45	8.71	0.73	5.90	0.33	4.49	-1.32	3.33
28<	288	-4.97	11.98	6.36	18.13	0.88	8.69	-2.75	5.37
All bonds	6,031	-0.33	5.90	1.50	8.31	0.42	7.90	-0.44	5.78

Table 3: **Summary Statistics for Pricing Errors on Coupon Bonds**
This table reports the mean pricing errors (Mean) and the root mean-squared pricing errors (RMSE) of the Canadian bond prices for the AFNS model with independent factors when evaluated at the filtered states. The AFNS model is estimated on three different data sets: (1) the universe of Canadian coupon bond prices, (2) zero-coupon yields constructed by the Bank of Canada, and (3) zero-coupon yields constructed from Canadian coupon bond prices using the Svensson (1995) discount function. The final two columns report the corresponding statistics when pricing coupon bonds using the estimated Svensson (1995) discount function. The pricing errors are reported in annual basis points and computed as the difference between the implied yield to maturity on the coupon bond and the model-implied yield to maturity on this bond. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

points for the Svensson (1995) yields. We emphasize that both the states and the model estimates in the AFNS model are here obtained from synthetic zero-coupon yields. Thus, the conventional two-step approach provides a fairly accurate fit to the underlying coupon bond prices, although they only enter indirectly in the estimated AFNS model through the synthetic zero-coupon yields. However, the one-step approach delivers an even better fit to these coupon bonds with an overall RMSE of only 5.90 basis points. Thus, going from the one-step approach to the two-step approach gives a deterioration in overall RMSE of 41% and 34% when using the Bank of Canada yields and the Svensson (1995) yields, respectively. This shows that the first step in the conventional two-step approach may add sizable noise to the predicted bond prices from an estimated DTSM.

These results are benchmarked in the final two columns of Table 3 to the fit of the Svensson (1995) discount function. That is, we compute the predicted coupon bond prices using the synthetic Svensson (1995) yields and express these pricing errors in yield to maturity. As

Maturity in months	Panel A: Bank of Canada yields				Panel B: Svensson (1995) yields			
	One-step approach		Two-step approach		One-step approach		Two-step approach	
	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
3	-4.80	20.14	0.06	9.77	-5.58	28.91	-3.17	9.67
6	-4.93	11.37	-0.95	3.52	-2.98	16.96	-0.11	1.95
12	-3.93	6.47	-1.20	9.00	-0.12	4.27	3.02	8.28
24	-0.14	5.02	1.34	8.64	0.98	5.74	3.19	8.66
36	1.30	4.75	2.39	6.64	0.18	4.26	0.80	4.33
60	0.26	4.07	1.11	3.71	-1.16	4.75	-2.99	6.98
84	-1.61	7.07	-1.27	6.76	-0.91	5.44	-3.42	7.39
120	-4.56	6.72	-5.39	8.72	0.82	4.73	-0.54	7.10
240	4.97	8.20	5.69	12.09	-0.15	6.04	4.77	11.64
360	-17.06	25.59	-2.05	8.24	-10.43	20.66	-2.06	10.26
All yields	-3.05	12.08	-0.03	8.11	-1.94	13.16	-0.05	8.09

Table 4: **Summary Statistics for Pricing Errors on Synthetic Zero-coupon Yields**
This table reports the mean pricing errors (Mean) and the root mean-squared pricing errors (RMSE) of the AFNS model with respect to synthetic zero-coupon yields from the Bank of Canada (in Panel A) and our implementation of Svensson (1995) yields (in Panel B). For the implementation of the two-step approach, we use Bank of Canada yields in Panel A and Svensson (1995) yields in Panel B of this table. All estimated versions of the AFNS model have independent factors and evaluate the fit at the filtered states. All numbers are measured in annual basis points. The data series are monthly covering the period from January 31, 2000, to April 30, 2016.

expected, the RMSEs for bonds with maturities exceeding two years are all smaller for the Svensson (1995) discount function than for any of the estimated AFNS models. However, the deterioration in fit for the estimated AFNS model based on the one-step approach is surprisingly small except for long-term bonds with more than 26 years to maturity. Indeed, for the zero to two-year maturity bucket, the one-step estimated AFNS model does *better* than the Svensson (1995) discount function (RMSE of 5.75 versus 8.83 basis points). Thus, for the construction of fitted yields, a one-step approach based on the AFNS model is about as accurate as fitting a flexible functional form like the Svensson (1995) curve to the data, which may be of independent interest to those wishing to construct a theoretically consistent synthetic zero-coupon yield curve.

When using the traditional two-step approach, the performance of DTSMs is normally evaluated by their ability to fit synthetic yields and not the underlying prices on coupon bonds. Therefore, we also consider the ability of the one-step estimated AFNS model to match the two samples of synthetic zero-coupon yields. Panel A in Table 4 evaluates fit to the Bank of Canada synthetic zero-coupon yields from the one-step estimated AFNS model and from the two-step AFNS model estimated using the Bank of Canada yields. As expected, the estimated AFNS model from the two-step approach has a much tighter fit to these synthetic yields than the estimated AFNS model from the one-step approach, notably in the short and long end of the yield curve. The second panel in Table 4 shows that we find the same pattern

when using the synthetic zero-coupon yields from the Svensson (1995) discount function. Thus, if synthetic zero-coupon yields are treated as "observed data," the econometrician would incorrectly prefer the estimated AFNS model from the two-step approach as the best representation of the Canadian bond market, although the estimated AFNS model from the one-step approach clearly provides the best fit to the observed coupon bond prices, as shown in Table 3.

3.4 Robustness of the One-Step Approach

The online Appendices D to H scrutinize the robustness of the one-step approach to four implementation choices. First, scaling the pricing errors by duration in equation (5) gives nearly identical results to representing coupon bond prices by their implied yield to maturity in the measurement equation. Omitting the duration-scaling of the pricing errors in equation (5) gives substantially different results, but we show in the online Appendix D that this alternative specification of the measurement equation is rejected by the data. Second, using the more accurate unscented Kalman filter (UKF) instead of the EKF to build the quasi log-likelihood function gives nearly identical results to those reported in Section 3.3. These results are also robust to using the sequential regression approach of Andreasen and Christensen (2015), which imposes fewer restrictions on the pricing errors and the state innovations than assumed in the EKF. Third, using a fully efficient ML estimator in the one-step approach is also shown to give nearly identical estimates to those based on QML and the EKF. The considered ML estimator is here derived by double limit asymptotics, where both the time series dimension and the number of bonds tend to infinity, as this allows us to obtain a likelihood function without resorting to simulation-based procedures. Finally, the one-step approach is also robust to considering weekly or even daily data, instead of the standard monthly data frequency adopted throughout the paper. Thus, the one-step approach with duration-scaled pricing errors and a QML estimator based on the EKF appears quite robust. However, the QML estimator based on the UKF or the proposed ML estimator could be helpful for estimating DTSMs with stronger deviations from linearity and Gaussianity than present in the AFNS model, for instance in models with stochastic volatility.

4 Simulation Study: Is One Step Better Than Two?

The preceding analysis has shown that the one- and two-step approaches give somewhat different estimates of DTSMs. But which approach gives the most accurate estimates? Here we answer this question by conducting a Monte Carlo study to analyze the finite-sample properties of estimating the AFNS model by the one- and two-step approaches. We first describe the formulation of the Monte Carlo study and then analyze the precision of the

estimated synthetic zero-coupon yields from applying a Svensson (1995) yield curve. The results for the estimated model parameters are reported in Section 4.3, while the accuracy of the filtered states and the standard yield curve decomposition are explored in Sections 4.4 and 4.5, respectively. Section 4.6 is devoted to the implementation of the two-step approach, where we explore how the number of synthetic yields and the adopted curve-fitting technique for these yields affect the DTSM estimates.

4.1 Setup for the Monte Carlo Study

Unlike previous simulation studies in the literature, our Monte Carlo study is formulated at the level of individual coupon bonds to account for estimation uncertainty in the construction of synthetic zero-coupon yields within the two-step approach. To get a representative data generating process for the Canadian bond market, we use the estimates of the AFNS model in the one-step approach from Table 2. Based on these parameters, we first simulate $N = 100$ samples for the three states at a monthly frequency for 196 months, which corresponds to the number of monthly observations in our Canadian sample.[11] These simulated sample paths for the states will be common across all exercises in the Monte Carlo study to facilitate the interpretation. The inputs for each of the two estimation approaches are then constructed as follows.

For the one-step approach, we use the simulated states to compute N panels of coupon-bond prices that match those observed in the Canadian sample in terms of available bonds and their characteristics. These bond prices are computed using the bond price formula in equation (4) in combination with the zero-coupon yields in equation (2). We then add measurement errors $\varepsilon_t^i \sim \mathcal{NID}\left(0, \sigma_\varepsilon^2\right)$ to the simulated bond prices and scale these errors by the duration of the simulated bond for consistency with equation (5).[12]

For the two-step approach, we take these simulated panels of coupon bond prices as input to extract synthetic zero-coupon yields based on the Svensson (1995) yield curve. For consistency with the empirical estimation results presented in the previous sections, we extract synthetic zero-coupon yields with ten constant maturities, 0.25, 0.5, 1, 2, 3, 5, 7, 10, 20, and 30 years, which we use for implementation of the two-step approach in the Monte Carlo study. Given that the underlying bond prices are already contaminated with measurement errors, we do not add additional noise to these synthetic yields.

To study the role of the data quality, we consider two cases where σ_ε is either 1 or 10 basis points. The first case with $\sigma_\varepsilon = 1$ basis point represents an ideal setting with hardly

[11] We simulate from (3) using a standard Euler-discretization, i.e., $X_t^i = X_{t-1}^i + \kappa_{ii}^{\mathbb{P}}(\theta_i^{\mathbb{P}} - X_{t-1}^i)\Delta t + \sigma_{ii}\sqrt{\Delta t} z_t^i$, where $z_t^i \sim \mathcal{N}(0,1)$ and $\Delta t = 0.0001$. The starting values X_0^i are drawn from the unconditional distribution of X_t.

[12] Note that we also use the same set of simulated samples of ε_t^i throughout the Monte Carlo study to make the results as comparable as possible.

Maturity	$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
in months	Mean	MAE	Mean	MAE
3	-4.09	6.79	-3.00	12.48
6	-2.37	3.89	-1.74	7.84
12	-0.32	1.05	-0.24	3.58
24	0.78	1.42	0.52	3.52
36	0.43	0.99	0.24	3.12
60	-0.48	1.05	-0.34	3.13
84	-0.67	1.26	-0.37	3.30
120	-0.34	0.94	-0.14	2.88
240	0.64	1.46	0.27	3.76
360	-1.32	3.09	-0.90	9.34

Table 5: **Accuracy of Estimated Svensson (1995) Yields**
The table reports the mean of the sampling distribution of the mean errors (Mean) and mean absolute errors (MAE) for each zero-coupon yield constructed using the Svensson (1995) discount function relative to the true zero-coupon yield implied by the AFNS model with simulated samples of length $T = 196$ and $N = 100$ repetitions. The mean is obtained by first computing the mean errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these means across the $N = 100$ simulated samples. Similarly, the MAE are obtained by first computing the mean absolute errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these absolute means across the $N = 100$ simulated samples. The true states are generated from the AFNS model as described in Section 4.1. All numbers are reported in annual basis points.

any noise in bond prices and helps to isolate the effects of the curve fitting procedure in the two-step approach. The second case with $\sigma_\varepsilon = 10$ basis points is included to describe a more realistic setting, as we find $\sigma_\varepsilon = 7$ basis points in our Canadian sample when using the one-step approach.

4.2 Accuracy of Synthetic Yields

We first consider the accuracy of the synthetic zero-coupon yields from the Svensson (1995) yield curve when estimated on the simulated coupon bond prices. That is, we compare the estimated synthetic yields to the *true* zero-coupon yields from the AFNS model without measurement errors.

With small measurement errors of $\sigma_\varepsilon = 1$ basis point, Table 5 shows that the mean errors are generally very close to zero within the one- to twenty-year maturity range but somewhat larger at the three- and six-month maturities (-4 and -2 basis points, respectively) and at the thirty-year maturity (-1.3 basis points). This means that the Svensson (1995) yield curve slightly overpredicts the level of the zero-coupon yields at the short and long end of the curve. The low mean absolute errors (MAE) of roughly 1 basis point show that yields within the

Par.	True value	One-step approach				Two-step approach			
		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
		Mean bias	Std. dev.	Mean bias	Std. dev.	Mean bias	Std. dev.	Mean bias	Std. dev.
$\kappa_{11}^{\mathbb{P}}$	**0.1060**	0.2734	0.2660	0.2951	0.2659	0.2009	0.2114	0.3549	0.2489
$\kappa_{22}^{\mathbb{P}}$	**0.2157**	0.2970	0.3022	0.2939	0.3002	0.3139	0.3069	0.4454	0.3883
$\kappa_{33}^{\mathbb{P}}$	**0.7255**	0.1801	0.3973	0.2043	0.3706	0.4264	0.6491	0.5756	0.5329
σ_{11}	**0.0052**	0.0000	0.0000	0.0000	0.0001	-0.0005	0.0003	-0.0004	0.0003
σ_{22}	**0.0103**	0.0000	0.0005	0.0000	0.0006	0.0002	0.0006	0.0014	0.0009
σ_{33}	**0.0207**	-0.0001	0.0007	0.0001	0.0014	0.0035	0.0037	0.0058	0.0020
$\theta_1^{\mathbb{P}}$	**0.0529**	0.0015	0.0088	-0.0019	0.0086	-0.0023	0.0086	-0.0016	0.0087
$\theta_2^{\mathbb{P}}$	**-0.0275**	0.0021	0.0108	0.0020	0.0108	0.0036	0.0112	0.0028	0.0110
$\theta_3^{\mathbb{P}}$	**-0.0230**	-0.0012	0.0098	-0.0006	0.0064	-0.0027	0.0062	-0.0031	0.0062
λ	**0.3747**	0.0000	0.0008	-0.0004	0.0056	0.0490	0.0314	0.0423	0.0264

Table 6: **Accuracy of the Parameter Estimates in the AFNS Model**
The table reports the mean estimate minus true value (Mean bias) and the standard deviation (Std. dev.) of the sampling distribution for each of the estimated parameters in the AFNS model when using QML in the one-step approach and ML in the two-step approach, where synthetic yields are generated with the Svensson (1995) yield curve, both with simulated samples of length $T = 196$ and $N = 100$ repetitions.

one- to twenty-year maturity spectrum are estimated very accurately, whereas yields at the short and long end of the curve are estimated less precisely (MAEs of 6.8 and 3.1 basis points for the three-month and thirty-year yields). This imprecision reflects the challenges of fitting the endpoints of a yield curve.

With larger measurement errors of $\sigma_\varepsilon = 10$ basis points, maturities between one and twenty years remain well approximated with mean errors close to zero. The precision of these yields in terms of the MAEs only decreases by a factor of three, which is substantially lower than the ten-fold increase in σ_ε. In most cases, the construction of the synthetic zero-coupon yields are able to correctly smooth out a large fraction of the idiosyncratic noise, ε_t^i, in the underlying bond prices, which generally leaves the measurement equation in the two-step approach with *smaller* errors than in the one-step approach. Accordingly, the synthetic zero-coupon yields from the Svensson (1995) yield curve appear to be very accurate in the present setting. Hence, for $\sigma_\varepsilon = 10$ basis points, the Monte Carlo results may favor the two-step approach, as the measurement equation here has reduced noise.

4.3 Estimated Parameters from One- and Two-step Approaches

Table 6 summarizes the outcome of the Monte Carlo study for the model parameters by reporting the mean and the standard deviation of each of the estimated coefficients in the AFNS model across simulations. We first note that both the one- and two-step approaches generate the familiar positive bias in the mean-reversion parameters $\left\{\kappa_{11}^{\mathbb{P}}, \kappa_{22}^{\mathbb{P}}, \kappa_{33}^{\mathbb{P}}\right\}$, as discussed in Bauer et al. (2012). For the persistence of the slope factor, $\kappa_{22}^{\mathbb{P}}$, and the curvature factor, $\kappa_{33}^{\mathbb{P}}$,

we find that these biases are somewhat larger for the two-step approach. For instance, with $\sigma_\varepsilon = 10$ basis points, the bias in $\kappa^{\mathbb{P}}_{22}$ is 0.29 in the one-step approach but 0.45 in the two-step approach. The results are mixed for the level factor, as the one-step estimate of $\kappa^{\mathbb{P}}_{11}$ has a smaller bias with $\sigma_\varepsilon = 10$ basis points but not with $\sigma_\varepsilon = 1$ basis points.

The estimates of the volatility parameters in Σ are basically unbiased in the one-step approach and estimated with great precision—both with small and large measurement errors. The corresponding estimates in the two-step approach display small biases with $\sigma_\varepsilon = 1$ basis points, which generally increase with larger measurement errors. All elements in $\theta^{\mathbb{P}}$ are generally close to their true values, although a careful inspection of Table 6 reveals that the biases in $\theta^{\mathbb{P}}$ typically are smaller in the one-step approach compared with the two-step approach.

The estimates of the decay parameter λ for the slope and curvature factor are centered exactly around its true value in the one-step approach and estimated with great precision—both with small and large measurement errors. For the two-step approach, we see small positive biases in the estimates of λ, and some relatively greater imprecision compared to the one-step approach. For instance, when $\sigma_\varepsilon = 10$ basis points, the standard deviation in the estimates of λ are 0.0264 in the two-step approach, but only 0.0056 in the one-step approach. These differences in the estimates of λ are of particular interest given the work of Björk and Christensen (1999), which shows that the static Nelson-Siegel and Svensson yield curves are inconsistent with no-arbitrage restrictions because the corresponding λ parameter(s) in these static models may change across time. The biased estimate of λ in the two-step approach implies that the curvature factor carries a greater weight on shorter-term yields and is less sensitive to longer-term yields relative to the true model. Since short-term yields are more volatile than long-term yields, this explains the positive bias in the two-step estimates of σ_{33}. Furthermore, as short-term yields also tend to be less persistent than long-term yields, this also explains the more severe upward bias in the two-step estimates of $\kappa^{\mathbb{P}}_{33}$.

4.4 Estimated States

For each simulated sample and its related set of estimated parameters, we next study the accuracy of the estimated states. Table 7 shows that the filtered states in the one-step approach are basically unbiased, as the mean errors with $\sigma_\varepsilon = 10$ basis points are 0.21, -0.14, and -0.58 basis points for the level, slope, and curvature factors, respectively. In contrast, the conventional two-step approach generates notable biases in the estimated states. Furthermore, these biases in the two-step approach are largely unrelated to the degree of noise in the bond prices, implying that these biases must originate from the use of the estimated synthetic zero-coupon yields.

To measure the efficiency of the filtered states, we compute the mean absolute errors in

State variable	One-step approach				Two-step approach			
	$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
	Mean	MAE	Mean	MAE	Mean	MAE	Mean	MAE
L_t	-0.41	1.16	0.21	5.51	-8.93	10.80	-3.05	9.66
S_t	0.42	1.16	-0.14	5.97	14.28	15.67	7.45	13.46
C_t	0.16	2.51	-0.58	19.07	-21.70	32.33	-25.06	42.62

Table 7: **Accuracy of Estimated States in the AFNS Model**
The table reports the mean errors (Mean) and mean absolute errors (MAE) of each estimated state variable in the AFNS model when using QML in the one-step approach and ML in the two-step approach, where synthetic yields are generated with the Svensson (1995) yield curve, both with simulated samples of length $T = 196$ and $N = 100$ repetitions. The mean error is obtained by first computing the mean errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these means across the $N = 100$ simulated samples. Similarly, the MAE are obtained by first computing the mean absolute errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these absolute means across the $N = 100$ simulated samples. The true states are generated from the AFNS model as described in Section 4.1. All numbers are reported in basis points.

each simulated sample of $T = 196$ observations, which we report in Table 7 by averaging across the $N = 100$ simulations. The states in the one-step approach are estimated very accurately with MAE of 1 to 2 basis points in the ideal case with $\sigma_\varepsilon = 1$ basis point. For the more realistic setting where $\sigma_\varepsilon = 10$ basis points, we find somewhat larger MAE of 6, 6, and 19 basis points for the level, slope, and curvature factors, respectively. The filtered states in the two-step approach are estimated much less accurately partly due to a lower number of cross-sectional observations to represent the yield curve compared with the one-step approach. Also, the efficiency of the state estimates in the two-step approach are much less affected by increasing the noise in the bond prices. For instance, the MAE of the level and slope factors are basically unaffected by the value of σ_ε. This feature of the two-step approach reflects the fact that the synthetic zero-coupon yields are able to smooth out much of the noise in the bond prices and mitigate the effects of measurement errors.

4.5 Accuracy of Yield Decomposition

DTSMs are often applied to decompose the yield curve into expected future short rates and term premiums. We next explore whether the use of the one-step approach improves the precision of this decomposition compared to the conventional two-step approach. Hence, let the τ-year term premium be defined as $TP_t(\tau) = y(\tau, X_t) - \frac{1}{\tau}\int_t^{t+\tau} E_t^{\mathbb{P}}[r_s]ds$, where $\frac{1}{\tau}\int_t^{t+\tau} E_t^{\mathbb{P}}[r_s]ds$ denotes expected future short rates, see formulas in the online Appendix I.

For each simulated sample and its related set of estimated parameters and states, we next decompose the yield curve into expected future short rates and term premiums in Table 8. The mean errors in expected future short rates (EXP) are somewhat closer to zero in

Component	One-step approach				Two-step approach			
	$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
	Mean	MAE	Mean	MAE	Mean	MAE	Mean	MAE
Two-year yield	0.09	0.24	0.08	2.04	0.66	1.23	0.77	2.80
Five-year yield	0.18	0.33	0.19	2.23	-0.21	0.88	-0.23	2.67
Ten-year yield	0.25	0.33	0.24	1.72	-0.57	1.40	-0.67	2.67
Two-year EXP	-1.96	24.09	-1.67	24.27	-7.71	24.60	-7.08	27.81
Five-year EXP	-3.76	42.51	-2.95	42.15	-9.85	41.25	-9.36	43.84
Ten-year EXP	-5.20	59.37	-3.64	58.49	-11.55	56.81	-11.02	58.67
Two-year TP	2.05	24.09	1.75	24.23	8.36	24.64	7.85	27.64
Five-year TP	3.95	42.52	3.13	42.33	9.64	41.36	9.13	44.16
Ten-year TP	5.44	59.38	3.88	58.53	10.98	56.85	10.34	58.72

Table 8: **Accuracy of the Yield Decomposition in the AFNS Model**
The table reports the mean errors (Mean) and mean absolute errors (MAE) for each component of the yield curve decomposition into expected future short rates (EXP) and term premium (TP) at various maturities. The mean error for a given maturity is obtained by first computing the mean errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these means across the $N = 100$ simulated samples. Similarly, the MAE for a given maturity is obtained by first computing the mean absolute errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these absolute means across the $N = 100$ simulated samples. All errors are shown in basis points and defined as the true value minus the model-implied value. The parameter and state estimates in the AFNS model are obtained by QML in the one-step approach and by ML in the two-step approach, where zero-coupon synthetic yields are generated with the Svensson (1995) yield curve. The true yields and expected future short rates are generated from the AFNS model as described in Section 4.1.

the one-step approach. This finding seems consistent with the smaller biases in $\mathcal{K}^{\mathbb{P}}$ and in the filtered states for the one-step approach reported in Section 4.3 and 4.4. The one-step approach also implies slightly lower mean errors for term premiums (TP) than the two-step approach. However, the MAE in Table 8 for expected future short rates and term premiums are very large and almost identical for both estimation approaches, meaning that the large estimation uncertainty clearly dominates the small improvement in mean errors for term premiums within the one-step approach.

Thus, standard yield curve decompositions do not benefit from the one-step approach. This is because the reported biases and large estimation uncertainty in term premiums originate from expected future short rates, and hence the estimated mean-reversion parameters in $\mathcal{K}^{\mathbb{P}}$, which are relatively insensitive to the number of cross-sectional observations used to represent the yield curve.

4.6 The Implementation of the Two-Step Approach

Given the widespread use of the conventional two-step approach, it seems useful to explore whether its performance can be improved compared to Sections 4.3 and 4.4. The essential

Par.	True value	Extended sample				Reduced sample			
		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
		Mean bias	Std. dev.	Mean bias	Std. dev.	Mean bias	Std. dev.	Mean bias	Std. dev.
$\kappa_{11}^{\mathbb{P}}$	**0.1060**	0.2549	0.2245	0.2854	0.2397	0.2742	0.2145	0.4359	0.3017
$\kappa_{22}^{\mathbb{P}}$	**0.2157**	0.2915	0.2961	0.3807	0.3518	0.3052	0.2959	0.3963	0.3488
$\kappa_{33}^{\mathbb{P}}$	**0.7255**	2.5730	2.3665	1.7932	1.6718	0.2011	0.4196	0.5085	0.5102
σ_{11}	**0.0052**	-0.0004	0.0004	-0.0001	0.0003	-0.0002	0.0004	0.0009	0.0004
σ_{22}	**0.0103**	0.0000	0.0005	0.0008	0.0007	0.0002	0.0005	0.0011	0.0007
σ_{33}	**0.0207**	0.0192	0.0135	0.0144	0.0084	-0.0012	0.0016	0.0028	0.0017
$\theta_1^{\mathbb{P}}$	**0.0529**	0.0027	0.0065	0.0021	0.0082	-0.0033	0.0084	-0.0021	0.0086
$\theta_2^{\mathbb{P}}$	**-0.0275**	-0.0015	0.0118	-0.0012	0.0115	0.0045	0.0107	0.0032	0.0108
$\theta_3^{\mathbb{P}}$	**-0.0230**	-0.0069	0.0097	-0.0048	0.0083	-0.0021	0.0060	-0.0030	0.0059
λ	**0.3747**	0.0026	0.0178	-0.0133	0.0210	0.0671	0.0308	0.0516	0.0256

Table 9: **Accuracy of the Parameter Estimates: The Number of Synthetic Yields**
The table reports the mean estimate minus true value (Mean bias) and the standard deviation (Std. dev.) of the sampling distribution for each of the estimated parameters in the AFNS model when using ML in the two-step approach, with simulated samples of length $T = 196$ and $N = 100$ repetitions. The extended samples consists of the 31 synthetic zero-coupon yields with maturities of 0.5, 1, 2, ..., 30 year from the Svensson (1995) yield curve. The reduced sample consists of six synthetic zero-coupon yields with maturities of 1, 2, 3, 5, 7, and 10 years from the Svensson (1995) yield curve.

decisions for the econometrician in the two-step approach are the number of synthetic zero-coupon yields to include and how to extract these yields from the panel of coupon bonds. We analyze how these choices affect the estimated DTSM.

4.6.1 The Number of Synthetic Yields

Our implementation of the two-step approach has so far used ten synthetic zero-coupon yields to represent the yield curve—a typical number in the literature. But this choice may affect the performance of the two-step approach. There may be a trade-off between bias and efficiency for the two-step estimated DTSM parameters. A large number of synthetic yields may increase efficiency but at the potential cost of including some less precisely measured yields, which could bias the estimated DTSM coefficients. Including only a few very accurately measured yields minimizes the risk of coefficient bias but at the potential cost of lower efficiency. We explore whether such a trade-off exists by implementing the two-step approach on two additional sets of simulated data samples that both are drawn from the *same* data generating process. The first is an extended sample of 31 synthetic yields with maturities of 0.5, 1, 2, ..., 30 year. The second is a reduced sample of only six yields with maturities of 1, 2, 3, 5, 7, and 10 years, where we omit the imprecisely estimated yields at the short and long end of the thirty-year yield curve.

The estimated coefficients in the AFNS model from this simulation exercise are reported in Table 9. We find somewhat surprisingly that the biases for $\kappa_{11}^{\mathbb{P}}$, $\kappa_{22}^{\mathbb{P}}$, σ_{11}, σ_{22}, and λ

State variable	Extended sample				Reduced sample			
	$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
	Mean	MAE	Mean	MAE	Mean	MAE	Mean	MAE
L_t	40.27	41.36	34.44	34.78	-19.59	20.33	-7.31	14.52
S_t	-35.83	37.46	-32.06	32.77	24.68	25.21	11.98	17.13
C_t	-65.82	67.28	-43.46	54.12	-14.52	31.40	-23.80	43.47

Table 10: **Accuracy of Estimated States: The Number of Synthetic Yields**
The table reports mean errors (Mean) and mean absolute errors (MAE) of each estimated state variable in the AFNS model using ML in the two-step approach. The mean error is obtained by first computing the mean errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these means across the $N = 100$ simulated samples. Similarly, the MAE are obtained by first computing the mean absolute errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these absolute means across the $N = 100$ simulated samples. The true states are generated from the AFNS model as described in Section 4.1. The extended samples consists of the 31 synthetic zero-coupon yields with maturities of 0.5, 1, 2, ..., 30 year from the Svensson (1995) yield curve. The reduced sample consists of six synthetic zero-coupon yields with maturities of 1, 2, 3, 5, 7, and 10 years from the Svensson (1995) yield curve. All numbers are reported in basis points.

are *smaller* in the extended sample compared to the standard sample (in Table 6) and the reduced sample. Still, it is hard to detect any efficiency gains from the extended sample for these parameters, except possibly for λ. Furthermore, we note that the estimates of $\kappa_{33}^{\mathbb{P}}$ and σ_{33} have the largest biases in the extended sample, which also provides the most imprecise estimates of $\kappa_{33}^{\mathbb{P}}$ and σ_{33}. The performance of the reduced sample with six synthetic yields is very similar to what we found for the standard sample with ten synthetic yields, meaning that the reduced sample avoids the large biases we occasionally find in the extended sample (e.g. in $\kappa_{33}^{\mathbb{P}}$ and σ_{33}). For each simulated sample and its related set of estimated parameters, Table 10 shows the accuracy of the filtered states. The extended sample has larger positive biases for the level factor and larger negative biases for the slope and curvature factor. The biases in the reduced sample are much smaller than in the extended sample, which explains the lower MAE of the filtered states for the reduced sample compared with the extended sample.

We draw two conclusions from this exercise. First, there does not appear to be an obvious trade-off between bias and efficiency in the two-step approach when varying the number of synthetic zero-coupon yields in the estimation of the DTSM. Second, the current practice of using a relatively low number of synthetic yields (i.e. between six and ten) seems well justified, at least when the synthetic yields are extracted based on the parametric discount function in Svensson (1995).

Par.	True value	Svensson (1995) yields				Nelson and Siegel (1987) yields			
		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
		Mean bias	Std. dev.	Mean bias	Std. dev.	Mean bias	Std. dev.	Mean bias	Std. dev.
$\kappa_{11}^{\mathbb{P}}$	**0.1060**	0.2397	0.2114	0.2489	0.2186	0.0260	0.1009	0.0551	0.1195
$\kappa_{22}^{\mathbb{P}}$	**0.2157**	0.3139	0.3069	0.4454	0.3883	0.3255	0.2813	0.4504	0.3739
$\kappa_{33}^{\mathbb{P}}$	**0.7255**	0.4264	0.6491	0.5756	0.5329	2.7574	2.9780	1.2863	1.1920
σ_{11}	**0.0052**	-0.0005	0.0003	-0.0004	0.0003	-0.0023	0.0003	-0.0019	0.0003
σ_{22}	**0.0103**	0.0002	0.0006	0.0014	0.0009	0.0004	0.0005	0.0015	0.0008
σ_{33}	**0.0207**	0.0035	0.0037	0.0058	0.0020	0.0210	0.0179	0.0125	0.0087
$\theta_1^{\mathbb{P}}$	**0.0529**	-0.0023	0.0086	-0.0016	0.0087	-0.0013	0.0073	-0.0027	0.0079
$\theta_2^{\mathbb{P}}$	**-0.0275**	0.0036	0.0112	0.0028	0.0110	0.0011	0.0127	0.0036	0.0114
$\theta_3^{\mathbb{P}}$	**-0.0230**	-0.0027	0.0062	-0.0031	0.0062	-0.0051	0.0110	-0.0028	0.0074
λ	**0.3747**	0.0490	0.0314	0.0423	0.0264	0.0694	0.0310	0.0778	0.0268

Table 11: **Accuracy of the Parameter Estimates: Different Synthetic Yields**
The table reports the mean estimate minus true value (Mean bias) and the standard deviation (Std. dev.) of the sampling distribution for each of the estimated parameters in the AFNS model when using ML in the two-step approach on synthetic yields from the Svensson (1995) and Nelson and Siegel (1987) yield curves. The true yields are generated from the AFNS model as described in Section 4.1, with simulated samples of length $T = 196$ and $N = 100$ repetitions. For both types of yields we use the same ten constant maturities, 0.25, 0.5, 1, 2, 3, 5, 7, 10, 20, and 30 years.

4.6.2 Synthetic Nelson and Siegel (1987) Yields

An obvious difference between the true zero-coupon yields from the AFNS model and those from the Svensson (1995) yield curve is that the latter allows for an extra "hump" at the long end of the yield curve compared to the AFNS model. Within our setting, this additional hump clearly seems redundant and we therefore explore the effects of omitting it when extracting synthetic zero-coupon yields from our simulated panels of bond prices. That is, we consider the case where the synthetic yields are constructed using the parametric discount function in Nelson and Siegel (1987), which is described further in online Appendix B. We continue to use the same ten constant yield maturities as in Section 4.3.

Table 11 reports the results for the estimated coefficients in the AFNS model from this simulation exercise, which we benchmark to the findings in Section 4.3 based on Svensson (1995) yields. We generally find that the estimated coefficients are adversely affected by using the more parsimonious specification of Nelson and Siegel (1987) to extract the synthetic yields. Most notably, the biases for $\kappa_{22}^{\mathbb{P}}$, σ_{11}, σ_{22}, σ_{33}, and λ increase somewhat when using the Nelson and Siegel (1987) yields compared to the Svensson (1995) yields, whereas the opposite applies for $\kappa_{11}^{\mathbb{P}}$. Table 12 further shows that the filtered state estimates with the Nelson and Siegel (1987) yields are less efficient (as measured by MAE) compared to the Svensson (1995) yields, whereas the mean errors are more similar across the two specifications.

To understand why the use of Nelson and Siegel (1987) yields worsen the performance of the two-step approach, recall that yields in the AFNS model include the convexity-adjustment

State variable	Svensson (1995) yields				Nelson and Siegel (1987) yields			
	$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points		$\sigma_\varepsilon = 1$ basis point		$\sigma_\varepsilon = 10$ basis points	
	Mean	MAE	Mean	MAE	Mean	MAE	Mean	MAE
L_t	-8.93	10.80	-3.05	9.66	6.69	39.44	-10.69	22.45
S_t	14.28	15.67	7.45	13.46	-2.78	37.30	15.13	26.69
C_t	-21.70	32.33	-25.06	42.62	-43.72	74.43	-24.83	56.51

Table 12: **Accuracy of Estimated States: Different Synthetic Yields**
The table reports the mean errors (Mean) and mean absolute errors (MAE) of each estimated state variable in the AFNS model using ML in the two-step approach based on Svensson (1995) and Nelson and Siegel (1987) yields, each with the same ten maturities, 0.25, 0.5, 1, 2, 3, 5, 7, 10, 20, and 30 years. The mean error is obtained by first computing the mean errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these means across the $N = 100$ simulated samples. Similarly, the MAE are obtained by first computing the mean absolute errors in each of the simulated samples across the $T = 196$ observations, and we then report the average of these absolute means across the $N = 100$ simulated samples. The true states are generated from the AFNS model as described in Section 4.1. All numbers are reported in basis points.

term $A(\tau)/\tau$ to ensure absence of arbitrage, which is not present in the Nelson and Siegel (1987) and Svensson (1995) specifications. This yield-adjustment term grows in size with maturity and becomes non-negligible for long-term bonds. Our results therefore suggest that the extra flexibility for long-term yields in the Svensson (1995) discount function is desirable in this context, because it allows us to capture this convexity-adjustment when estimating synthetic zero-coupon yields. Thus, restricting the parametric discount function for extracting synthetic zero-coupon yields does not improve performance, which supports the wide use of the synthetic U.S. Treasury yields constructed by Gürkaynak et al. (2007, 2010).

4.7 Summarizing the Main Insights From the Monte Carlo Study

The main insight from this Monte Carlo study is that a DTSM may be estimated more reliably by using directly observed market prices on coupon bonds instead of synthetic zero-coupon yields. Although these synthetic zero-coupon yields are estimated very accurately with well-established curve-fitting techniques, we nevertheless find that seemingly negligible errors in these synthetic yields do affect the estimated parameters in a DTSM. In particular, all risk-neutral parameters are estimated with smaller biases and greater efficiency with a one-step approach compared with the conventional two-step approach. In large part, this improvement reflects the denser representation of the yield curve used in the one-step approach as well as the avoidance of estimation errors introduced by synthetic zero-coupon yields. We also find that parameters in the $\mathbb{P}$-dynamics benefit from a one-step approach, although these parameters are unrelated to the $\mathbb{Q}$-dynamics with an essentially affine specification for the market prices of risk. This improvement therefore arises mainly because the states are estimated with lower

Par.	One-step approach		Two-step approach			
			Bank of Canada yields		Svensson (1995) yields	
	Est	SE	Est	SE	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.1450	0.0891	0.1373	0.0669	0.0521	0.0468
$\kappa_{22}^{\mathbb{P}}$	0.1066	0.0819	0.1074	0.0701	0.5166	0.1828
$\kappa_{33}^{\mathbb{P}}$	0.4337	0.2897	0.3503	0.3573	0.2646	0.3935
σ_{11}	0.0073	0.0002	0.0083	0.0003	0.0059	0.0002
σ_{22}	0.0122	0.0011	0.0098	0.0008	0.0114	0.0011
σ_{33}	0.0180	0.0021	0.0205	0.0021	0.0220	0.0026
$\theta_1^{\mathbb{P}}$	0.0546	0.0034	0.0546	0.0115	0.0555	0.0242
$\theta_2^{\mathbb{P}}$	-0.0396	0.0087	-0.0248	0.0135	-0.0231	0.0046
$\theta_3^{\mathbb{P}}$	-0.0248	0.0117	-0.0250	0.0135	-0.0238	0.0219
λ	0.3920	0.0123	0.3473	0.0135	0.4754	0.0149

Table 13: **Estimated Parameters in the B-AFNS Model**
This table reports the estimated parameters (Est) in the B-AFNS model with independent factors and their standard errors (SE) using either the one-step or the two-step approach. The SE are in all cases computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, computed as outlined in Harvey (1989). The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

biases and greater precision in the one-step approach.[13]

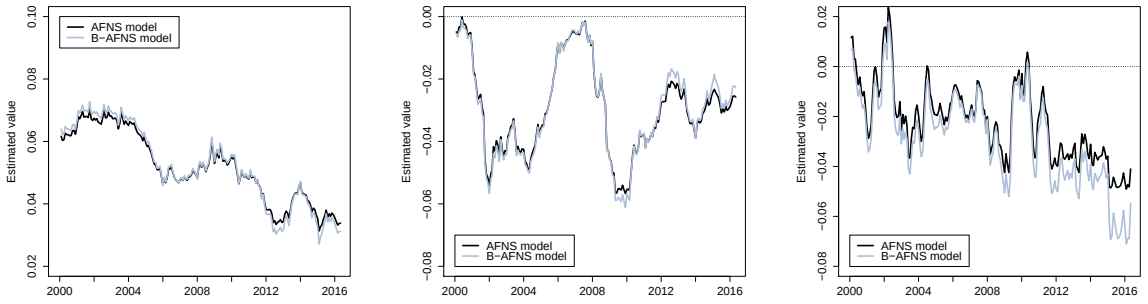
5 Forecasting Bond Yields

The previous analysis has shown that parameters and states in the AFNS model are estimated with smaller biases and greater efficiency by the one-step approach when compared to the two-step approach. The present section explores whether these advantages are sufficiently large to improve the ability of the AFNS model to forecast yields out of sample. To make this forecasting exercise more comprehensive, we include two additional models: a shadow-rate specification to accommodate the zero lower bond and a five-factor AFNS model to better fit long-term bonds. From a methodological perspective, these extensions also illustrate that the one-step approach is applicable to nonlinear DTSMs and to models with more than three states. As in the previous section, we benchmark the performance of the one-step approach to those from the two-step approach based on synthetic zero-coupon yields.

5.1 A Shadow-Rate Model

Given the very low policy rates in many economies during the recent financial crisis and its aftermath, it has become popular to account for the zero lower bound (ZLB) in DTSMs.

[13] Our current Monte Carlo study uses a small number of simulations ($N = 100$), unreported results using $N = 1{,}000$ suggest that little changes by increasing N.



(a) L_t (b) S_t (c) C_t

Figure 3: **Estimated States in the One-Step Approach**
This figure reports the filtered estimates of level, slope, and curvature in the AFNS and B-AFNS model. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

Although relative to other developed economies, Canadian short rates were close to zero for only a limited period in our sample (as seen from Figure 2(a)), it is still possible that the ZLB may affect the shape and dynamics of the yield curve. To enforce the ZLB in the AFNS model, we introduce the shadow rate $s_t = L_t + S_t$ and let $r_t = \max\{0, s_t\}$, as in Christensen and Rudebusch (2015). All other aspects of this B-AFNS model remain as described above for the AFNS model.[14] The expression for zero-coupon yields in the B-AFNS model is not available in closed form but approximated numerically using the accurate method of Krippner (2013).[15]

Table 13 shows that all elements in $\mathcal{K}^{\mathbb{P}}$ and $\theta^{\mathbb{P}}$ in the B-AFNS model are estimated very imprecisely across the three data sets, which is similar to our finding for the AFNS model. The volatility parameters in Σ are estimated much more precisely and are generally higher in the B-AFNS model when compared to the AFNS model. Figure 3 shows that this difference is mainly explained by greater variability of the states after 2008, because the shadow-rate specification in the B-AFNS model allows the states to move more freely than seen in the AFNS model without violating the ZLB. We also find that λ is estimated to be somewhat higher in all three data sets when accounting for the ZLB in comparison to the AFNS model. Similar to the pattern observed for the AFNS model, the estimate of λ in the B-AFNS model using the one-step approach lies in between those from the two-step approach, as λ is 0.392 in the one-step approach, 0.347 in the two-step approach based on Bank of Canada yields, and 0.475 in the two-step approach using Svensson (1995) yields.

Table 14 reports the pricing errors of the B-AFNS model for the underlying coupon bonds. For the one-step approach and both versions of the two-step approach, we find slightly smaller

[14] Following Kim and Singleton (2012), the prefix "B-" refers to a shadow-rate model in the spirit of Black (1995).

[15] See also Christensen and Rudebusch (2015, 2016) for further details on this approximation and its accuracy.

Maturity bucket in years	No. obs.	One-step approach		Two-step approach			
				Bank of Canada yields		Svensson (1995) yields	
		Mean	RMSE	Mean	RMSE	Mean	RMSE
0-2	1,472	-0.41	5.51	2.91	7.84	0.70	9.54
2-4	1,098	0.47	4.92	1.55	6.92	1.26	5.20
4-6	744	-0.36	4.17	0.14	4.67	-2.06	5.08
6-8	404	-1.79	5.82	-1.04	5.55	-3.62	7.07
8-10	477	-3.65	6.75	-3.04	6.89	-4.76	7.77
10-12	289	-2.35	6.80	-1.23	8.55	-1.80	8.02
12-14	155	2.54	5.62	4.93	11.28	5.12	10.55
14-16	168	-0.34	4.17	-0.96	8.66	0.89	7.90
16-18	179	-0.30	4.79	-1.68	9.93	1.13	8.80
18-20	192	0.94	3.93	1.89	8.33	4.66	8.26
20-22	186	3.31	5.38	5.44	9.58	7.43	10.07
22-24	142	1.49	5.38	2.90	6.82	4.74	6.89
24-26	124	1.45	5.40	4.16	8.20	4.78	7.37
26-28	113	-2.98	6.88	1.14	5.66	0.43	4.45
28<	288	-2.59	9.24	1.16	5.95	-1.71	4.59
All yields	6,031	-0.52	5.62	1.15	7.34	0.15	7.57

Table 14: **Summary Statistics of Bond Fitted Errors in the B-AFNS Model**
This table reports the mean pricing errors (Mean) and the root mean-squared pricing errors (RMSE) of the Canadian bond prices for the B-AFNS model with independent factors. The pricing errors are reported in annual basis points and computed as the difference between the implied yield on the coupon bond and the model-implied yield on this bond. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

RMSEs in the B-AFNS model compared to the AFNS model. For instance, the overall RMSE falls by 5% from 5.90 to 5.62 basis points in the one-step approach. Thus, accounting for the ZLB does not materially improve the ability of the AFNS model to match Canadian coupon bond prices.

5.2 A Five-Factor Model

The main motivation of Gürkaynak et al. (2007) to prefer the Svensson (1995) curve over the simpler specification in Nelson and Siegel (1987) is that the Svensson (1995) curve allows for an additional hump that helps fit U.S. bond yields beyond the ten- to fifteen-year maturity spectrum. The AFNS model may potentially also benefit from additional dynamics to fit long-term Canadian bond prices, as its loadings for the slope and curvature factor decay to zero as maturity approaches infinity. This often implies (for reasonable values of λ) that only the level factor in the AFNS model can be used to fit long-term bonds, which may at times be insufficient as noted in Christensen et al. (2011).

To explore whether the performance of the AFNS model on our Canadian sample may

be improved further, we consider the generalized AFNS model of Christensen et al. (2009), which includes an additional pair of slope and curvature factors that help to fit long-term bonds. In this AFGNS model $r_t = L_t + S_t + \tilde{S}_t$, where $\tilde{S}_t$ is an additional (long-term) slope factor. The state dynamics under the risk-neutral $\mathbb{Q}$-measure is

$$\begin{pmatrix} dL_t \\ dS_t \\ d\tilde{S}_t \\ dC_t \\ d\tilde{C}_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & -\lambda & 0 \\ 0 & 0 & \tilde{\lambda} & 0 & -\tilde{\lambda} \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \tilde{\lambda} \end{pmatrix} \left[\begin{pmatrix} \theta_1^{\mathbb{Q}} \\ \theta_2^{\mathbb{Q}} \\ \theta_3^{\mathbb{Q}} \\ \theta_4^{\mathbb{Q}} \\ \theta_5^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ \tilde{S}_t \\ C_t \\ \tilde{C}_t \end{pmatrix} \right] dt + \Sigma d\tilde{W}_t^{\mathbb{Q}},$$

where $\lambda > \tilde{\lambda} > 0$ and $\tilde{C}_t$ is an additional (long-term) curvature factor. Zero-coupon yields are then given by

$$\begin{aligned} y(t,T) &= L_t + \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} S_t + \left[\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)} \right] C_t \\ &\quad + \frac{1-e^{-\tilde{\lambda}(T-t)}}{\tilde{\lambda}(T-t)} \tilde{S}_t + \left[\frac{1-e^{-\tilde{\lambda}(T-t)}}{\tilde{\lambda}(T-t)} - e^{-\tilde{\lambda}(T-t)} \right] \tilde{C}_t - \frac{\tilde{A}(t,T)}{T-t}, \end{aligned}$$

where the yield-adjustment term $\tilde{A}(t,T)$ is derived in Christensen et al. (2009). The $\mathbb{P}$-dynamics for this five-factor model is obtained in a standard fashion by adopting an essential affine specification for the market price of risk, as in Section 3.1.

The estimation results for the AFGNS model are reported in Table 15—again, limited to independent factor dynamics under the $\mathbb{P}$-measure. The decay parameter λ is estimated to be somewhat larger than in the AFNS model, because S_t and C_t no longer have to fit long-term bonds. The very low estimate of the second decay parameter $\tilde{\lambda}$ implies that the additional factors $\tilde{S}_t$ and $\tilde{C}_t$ greatly assist the level factor in matching the long end of the Canadian yield curve.

Table 16 reports the pricing errors of the AFGNS model for the underlying coupon bonds, and we clearly see that all three versions of this five-factor model provide a closer fit to nearly all bonds when compared with the AFNS model. This is highligthed in Table 16 by ΔRMSE, which shows the difference in RMSE between the AFGNS model and the AFNS model for the one-step approach and for the two implementations of the two-step approach. For the one-step approach, we see large improvements in the RMSEs for long-term bonds, but also in the zero to two-year and two- to four-year maturity buckets, which both contain a large number of bonds. As a result, the overall RMSE within the one-step approach drops from 5.90 basis points in the AFNS model to just 4.51 basis points in this extended model, which corresponds to a 24% reduction in the size of the in-sample fitted errors. This also means that

Par.	One-step approach		Two-step approach			
			Bank of Canada yields		Svensson (1995) yields	
	Est	SE	Est	SE	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.0453	0.0484	0.3656	0.4340	0.1365	0.1194
$\kappa_{22}^{\mathbb{P}}$	0.1835	0.1418	0.6233	0.3229	0.5279	0.2508
$\kappa_{33}^{\mathbb{P}}$	0.2015	0.1973	0.1214	0.1790	1.1160	0.3580
$\kappa_{44}^{\mathbb{P}}$	0.7371	0.3144	0.9582	0.4657	1.1599	0.3570
$\kappa_{55}^{\mathbb{P}}$	0.1970	0.1255	0.3950	0.2917	0.1464	0.1347
σ_{11}	0.0031	0.0006	0.0092	0.0007	0.0050	0.0002
σ_{22}	0.0125	0.0010	0.0125	0.0009	0.0140	0.0009
σ_{33}	0.0106	0.0009	0.0093	0.0009	0.0150	0.0015
σ_{44}	0.0237	0.0019	0.0209	0.0013	0.0359	0.0021
σ_{55}	0.0200	0.0014	0.0215	0.0033	0.0189	0.0013
$\theta_1^{\mathbb{P}}$	0.0500	0.0053	0.0982	0.0110	0.0580	0.0091
$\theta_2^{\mathbb{P}}$	0.0183	0.0172	-0.0040	0.0087	0.0069	0.0099
$\theta_3^{\mathbb{P}}$	-0.0452	0.0140	-0.0691	0.0219	-0.0456	0.0045
$\theta_4^{\mathbb{P}}$	0.0064	0.0090	-0.0041	0.0071	0.0061	0.0093
$\theta_5^{\mathbb{P}}$	0.0486	0.0185	-0.0318	0.0229	0.0252	0.0329
λ	0.6416	0.0280	1.3699	0.0297	0.9290	0.0102
$\tilde{\lambda}$	0.1166	0.0084	0.0786	0.0039	0.1185	0.0026

Table 15: **Parameter Estimates in the AFGNS Model**
This table reports the estimated parameters (Est) in the AFGNS model with independent factors and their standard errors (SE) using either the one-step or the two-step approach. The SE in the one-step approach are computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, computed as outlined in Harvey (1989). The SE in the two-step approach are computed from the inverse of the variance of the score. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

the AFGNS model clearly provides a *better* overall fit to the bond prices than the Svensson (1995) discount function with an overall RMSE of 5.78 basis points. Given this satisfying performance of the AFGNS model, its zero-coupon yields may thus be used as another and slightly more accurate representation of the Canadian yield curve than the zero-coupon yields from the Svensson (1995) discount function.

5.3 Forecast Exercise and Results

We structure the forecast exercise to match the Consensus Forecasts, which is a monthly survey of professional forecasters. Those survey participants submit forecasts of the three-month Treasury bill rate and the ten-year government bond yield for the next three and twelve months. To measure the realization of the three-month Treasury bill rate, we linearly interpolate between two Treasury bills whose remaining times to maturity constitute the tightest bracket around the three-month maturity point. For the ten-year government yield,

Maturity bucket in years	No. obs.	One-step approach			Two-step approach					
					Bank of Canada yields			Svensson (1995) yields		
		Mean	RMSE	ΔRMSE	Mean	RMSE	ΔRMSE	Mean	RMSE	ΔRMSE
0-2	1,472	-0.20	4.86	-0.88	2.63	5.56	-2.31	-0.73	9.01	-0.80
2-4	1,098	0.36	3.82	-0.92	-0.93	4.49	-2.71	0.33	4.28	-1.33
4-6	744	0.77	3.77	-0.08	-0.03	3.29	-1.08	0.06	3.91	-0.70
6-8	404	-0.78	3.88	-1.59	1.54	3.96	-1.06	-1.47	4.32	-2.29
8-10	477	-2.39	5.33	-0.74	0.91	5.00	-1.61	-2.28	5.31	-2.50
10-12	289	-2.01	6.45	0.25	2.14	7.98	-0.58	-1.68	6.00	-2.35
12-14	155	1.87	4.01	-2.71	7.45	10.56	-1.42	2.59	4.86	-6.32
14-16	168	0.03	2.45	-1.87	1.71	6.36	-2.69	0.61	3.04	-5.24
16-18	179	-0.21	3.10	-1.56	-1.62	7.70	-2.19	-0.06	4.13	-4.71
18-20	192	0.71	3.77	-0.56	-0.53	5.92	-2.97	0.44	4.41	-4.23
20-22	186	2.39	5.10	0.13	1.17	4.74	-5.32	2.47	4.69	-5.63
22-24	142	1.82	5.31	0.84	-0.92	4.51	-2.76	0.99	3.67	-3.42
24-26	124	1.26	4.48	-0.53	0.16	5.54	-2.83	1.44	3.66	-3.77
26-28	113	-1.21	4.51	-3.85	-1.51	4.71	-1.19	-1.56	3.51	-0.98
28<	288	-1.88	5.32	-6.59	11.59	32.98	14.85	0.33	7.82	-0.87
All bonds	6,031	-0.13	4.51	-1.28	1.46	8.93	0.62	-0.26	6.05	-1.85

Table 16: **Summary Statistics of Bond Fitted Errors in the AFGNS Model**
This table reports the mean pricing errors (Mean) and the root mean-squared pricing errors (RMSE) of Canadian coupon bond prices for the AFGNS model with independent factors. The table also reports the difference in RMSE (ΔRMSE) between the AFGNS model and the AFNS model within the one-step approach and each of the two implementations of the two-step approach. All pricing errors are reported in annual basis points and computed as the difference between the implied yield on the coupon bond and the model-implied yield on this bond. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

we first note that the Bank of Canada (like the U.S. Treasury) tends to issue new bonds as close to par as possible, and we therefore interpret the survey participants as forecasting the ten-year par-coupon yield. The Bank of Canada only issues new ten-year bonds roughly once a year, and we therefore exploit our finding in Section 5.2 and use an estimated version of the AFGNS model to compute accurate realizations of the ten-year par yield. All three DTSMs in this forecasting study are estimated by the one-step approach and by the two-step approach using synthetic zero-coupon yields from the Bank of Canada, the Svensson (1995) discount function, and the Nelson and Siegel (1987) yield curve. To get a reasonable handle on the persistence of the states in the three models, we begin the forecast analysis in December 2006. Further details related to the implementation of the forecast study are provided in the online Appendix J.

Starting with the three-month ahead forecasts of the three-month yield in Table 17, the one-step approach does clearly better than the two-step approach using Svensson (1995) yields and Nelson and Siegel (1987) yields for all three models. However, we see the opposite pattern when the two-step approach is implemented using the Bank of Canada yields. The one-step

Model	Three-month forecasts			Twelve-month forecasts		
	Mean	RMSE	MAE	Mean	RMSE	MAE
Consensus Forecasts	-18.47	41.50	24.18	-84.97	122.25	85.07
AFNS model:						
One-step approach	-34.50	55.27	39.68	-79.52	115.80	89.08
Two-step approach						
Bank of Canada yields	-33.36	52.52	36.57	-91.75	126.53	96.11
Svensson (1995) yields	-37.15	56.87	43.32	-83.10	115.90	88.01
Nelson and Siegel (1987) yields	-43.76	63.02	47.74	-81.90	116.92	84.79
B-AFNS model:						
One-step approach	-36.87	55.80	40.03	-85.34	117.83	91.56
Two-step approach						
Bank of Canada yields	-33.52	52.63	36.59	-90.66	128.83	96.17
Svensson (1995) yields	-42.66	58.49	44.95	-96.84	121.68	96.95
Nelson and Siegel (1987) yields	-55.96	71.21	58.50	-108.40	135.47	109.94
AFGNS model:						
One-step approach	-30.46	51.44	33.96	-72.42	107.72	77.87
Two-step approach						
Bank of Canada yields	-20.71	43.71	28.83	-61.21	88.19	64.12
Svensson (1995) yields	-48.28	68.40	55.72	-106.33	131.49	109.39
Nelson and Siegel (1987) yields	-56.85	75.06	61.04	-105.63	129.74	105.63

Table 17: **Summary Statistics of Three-Month Yield Forecast Errors**
This table reports the mean forecasting errors (Mean), the root mean squared forecasting errors (RMSE), and the mean absolute forecasting errors (MAE). All forecasts are computed from DTSMs that are estimated recursively with diagonal matrices for $\mathcal{K}^P$ and Σ. The forecast errors are reported as the true value minus the model-implied prediction, and all numbers are reported in annual basis points.

approach continues to do well for the three-month yield when forecasting twelve months ahead, and it delivers the most accurate predictions in the AFNS and B-AFNS model. We also find that the forecasts from the AFNS model are generally not improved by accounting for the ZLB via the B-AFNS model, which is likely explained by the relatively brief period that the Canadian short rate stayed at the ZLB during our sample. Instead, the forecasts from the AFNS model are greatly improved by using the more flexible AFGNS model, but only when using the one-step approach and the two-step approach based on Bank of Canada yields,

Model	Three-month forecasts			Twelve-month forecasts		
	Mean	RMSE	MAE	Mean	RMSE	MAE
Consensus Forecasts	-15.00	47.34	38.62	-78.69	101.77	87.44
AFNS model:						
One-step approach	-24.29	43.26	35.45	-70.27	86.41	72.85
Two-step approach						
Bank of Canada yields,	-29.34	46.88	38.54	-86.63	100.09	87.73
Svensson (1995) yields	-23.68	43.00	35.12	-68.94	86.21	73.51
Nelson and Siegel (1987) yields	-21.34	41.02	33.20	-60.05	79.39	67.83
B-AFNS model:						
One-step approach	-22.84	42.50	34.46	-67.65	84.65	71.19
Two-step approach						
Bank of Canada yields	-26.56	45.40	37.11	-81.41	97.07	84.45
Svensson (1995) yields	-22.70	42.62	34.32	-68.92	88.57	75.49
Nelson and Siegel (1987) yields	-23.57	43.71	34.83	-67.09	89.09	74.97
AFGNS model:						
One-step approach	-19.04	41.86	34.52	-62.34	81.71	68.57
Two-step approach						
Bank of Canada yields	-14.02	41.42	33.33	-58.80	84.23	67.50
Svensson (1995) yields	-29.68	47.78	39.33	-82.96	96.96	84.49
Nelson and Siegel (1987) yields	-28.44	45.40	37.32	-74.73	88.73	76.32

Table 18: **Summary Statistics of Ten-Year Yield Forecast Errors**
This table reports the mean forecasting errors (Mean), the root mean squared forecasting errors (RMSE), and the mean absolute forecasting errors (MAE). All forecasts are computed from DTSMs that are estimated recursively with diagonal matrices for $\mathcal{K}^P$ and Σ. The forecast errors are reported as the true value minus the model-implied prediction, and all numbers are reported in annual basis points.

which both outperform the Consensus Forecasts when forecasting twelve months ahead.

The corresponding forecasts for the ten-year bond yield are summarized in Table 18. We once again find that the one-step approach delivers competitive forecasts when compared to the two-step approach, where the performance depends crucially on the chosen set of synthetic yields. That is, for the AFNS and B-AFNS model, we obtain the best results in the two-step approach by using the Nelson and Siegel (1987) yields, whereas we get the best results for the AFGNS model when using Bank of Canada yields. The one-step approach is in this sense

more robust, as it either gives the best forecasts or is close to the best performing approach for all three models. We also note that the flexible five-factor AFGNS model also delivers the best forecasts for the ten-year bond yield among the three DTSMs, but only when using the one-step approach and the two-step approach based on Bank of Canada yields.

Thus, the one-step approach can also be used to generate competitive out-of-sample forecasts when compared to surveys and the conventional two-step approach.

6 Conclusion

This paper demonstrates the advantages of estimating DTSMs directly on actual bond prices as opposed to the usual approach of using a limited number of synthetic zero-coupon yields. For our empirical sample of Canadian bonds, we find that seemingly small differences between two data sets of synthetic yields can affect the estimated DTSM parameters. Furthermore, we find that a one-step DTSM estimation gives a substantially closer fit to actual bond prices than a two-step approach. Accordingly, the use of synthetic yields in the conventional two-step approach may add some non-negligible noise to the predicted bond prices from an estimated DTSM.

We also explore the finite-sample properties of the one- and two-step approaches in a Monte Carlo study. A novel feature of this simulation exercise is that it is formulated at the level of individual coupon bond prices, so we can account for estimation uncertainty in the construction of the synthetic zero-coupon yields. A key insight from this Monte Carlo study is that the risk-neutral parameters in DTSMs are estimated with smaller biases and greater efficiency in the one-step approach when compared to the two-step approach.

There are likely additional advantages to estimating DTSMs directly on bond prices. For example, DTSMs could be augmented with bond-specific characteristics to assess liquidity premiums as in Fontaine and Garcia (2012) and Andreasen et al. (2018). Similarly, one could expand the work of Pancost (2018) and use the one-step approach to determine the individual bonds that trade cheap or at a premium relative to the overall market as use this information for portfolio management, arbitrage trading, or market surveillance. We leave these and other applications for future research.

References

Andreasen, Martin M. and Bent Jesper Christensen, 2015, "The SR Approach: A New Estimation Procedure for Non-Linear and Non-Gaussian Dynamic Term Structure Models," *Journal of Econometrics*, Vol. 184, No. 2, 420-451.

Andreasen, Martin M., Jens H. E. Christensen, and Simon Riddell, 2018, "The TIPS Liquidity Premium," Working Paper 2017-11, Federal Reserve Bank of San Francisco.

Bauer, Michael D., Glenn D. Rudebusch, and Jing (Cynthia) Wu, 2012, "Correcting Estimation Bias in Dynamic Term Structure Models," *Journal of Business and Economic Statistics*, Vol. 30, No. 3, 454-467.

Björk, Tomas and Bent J. Christensen, 1999, "Interest Rate Dynamics and Consistent Forward Rate Curves," *Mathematical Finance*, Vol. 9, 323-348.

Black, Fisher, 1995, "Interest Rates as Options," *Journal of Finance*, Vol. 50, No. 7, 1371-1376.

Bliss, Robert R., 1997, "Testing Term Structure Estimation Methods," *Advances in Futures and Options Research*, Vol. 9, 197-231.

Bolder, David J., Grahame Johnson, and Adam Metzler, 2004, "An Empirical Analysis of the Canadian Term Structure of Zero-Coupon Interest Rates," Bank of Canada Working Paper No. 2004-48.

Campbell, John Y. and Robert J. Shiller, 1991, "Yield Spread and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, Vol. 58, No. 3 , 495-514.

Christensen, Jens H. E., Francis X. Diebold, and Glenn D. Rudebusch, 2009, "An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model," *Econometrics Journal*, Vol. 12, No. 3, C33-C64.

Christensen, Jens H. E., Francis X. Diebold, and Glenn D. Rudebusch, 2011, "The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models," *Journal of Econometrics*, Vol. 164, No. 1, 4-20.

Christensen, Jens H. E. and Glenn D. Rudebusch, 2015, "Estimating Shadow-Rate Term Structure Models with Near-Zero Yields," *Journal of Financial Econometrics*, Vol. 13, No. 2, 226-259.

Christensen, Jens H. E. and Glenn D. Rudebusch, 2016, "Modeling Yields at the Zero Lower Bound: Are Shadow Rates the Solution?," in Eric Hillebrand, Siem Jan Koopman

(ed.) *Dynamic Factor Models (Advances in Econometrics, Volume 35)* Emerald Group Publishing Limited, 75-125.

Dai, Qiang and Kenneth J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, Vol. 55, No. 5, 1943-1978.

Dai, Qiang, Kenneth J. Singleton, and Wei Yang, 2004, "Predictability of Bond Risk Premia and Affine Term Structure Models," Manuscript, Graduate School of Business, Stanford University.

Diez de los Rios, Antonio, 2015, "A New Linear Estimator for Gaussian Dynamic Term Sructure Models," *Journal of Business and Economic Statistics*, Vol. 33, No. 1, 282-295.

Duffee, Gregory R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, Vol. 57, No. 1, 405-443.

Fama, Eugene F., 1976, "Forward Rates as Predictors of Future Spot Rates," *Journal of Financial Economics*, Vol. 3, No. 4, 361-377.

Faria, Adriano and Caio Almeida, 2017, "A Hybrid Spline-Based Parametric Model for the Real Yield Curve," Manuscript, FGV Rio de Janeiro.

Filipović, Damir and Sander Willems, 2016, "Exact Smooth Term Structure Estimation," Manuscript, EPFL and Swiss Finance Institute.

Fontaine, Jean-Sébastien and René Garcia, 2012, "Bond Liquidity Premia," *Review of Financial Studies*, Vol. 25, No. 4, 1207-1254.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2007, "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, Vol. 54, No. 8, 2291-2304.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2010, "The TIPS Yield Curve and Inflation Compensation," *American Economic Journal: Macroeconomics*, Vol. 2, No. 1, 70-92.

Harvey, A.C., 1989, *Forecasting, structural time series models and the Kalman filter,* Cambridge: Cambridge University Press.

Hu, Grace Xing, Jun Pan, and Jiang Wang, 2013, "Noise as Information for Illiquidity," *Journal of Finance*, Vol. 68, No. 6, 2341-2382.

Kim, Don H. and Kenneth J. Singleton, 2012, "Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields," *Journal of Econometrics*, Vol. 170, No. 1, 32-49.

Krippner, Leo, 2013, "A Tractable Framework for Zero Lower Bound Gaussian Term Structure Models," Discussion Paper 2013-02, Reserve Bank of New Zealand.

Nelson, Charles R. and Andrew F. Siegel, 1987, "Parsimonious Modeling of Yield Curves," *Journal of Business*, Vol. 60, No. 4, 473-489.

Pancost, N. Aaron, 2018, "Zero-Coupon Yields and the Cross-Section of Bond Prices," *Chicago University, Working Paper*, 1-66.

Steeley, James M., 2008, "Testing Term Structure Estimation Methods: Evidence from the UK STRIPs Market," *Journal of Money, Credit and Banking*, Vol. 40, No. 7, 1489-1512.

Svensson, Lars E. O., 1995, "Estimating Forward Interest Rates with the Extended Nelson-Siegel Method," *Quarterly Review*, No. 3, Sveriges Riksbank, 13-26.

Online Appendix

Term Structure Analysis with Big Data: One-Step Estimation Using Bond Prices

Martin M. Andreasen
Aarhus University, CREATES, and the Danish Finance Institute
mandreasen@econ.au.dk

Jens H. E. Christensen
Federal Reserve Bank of San Francisco
jens.christensen@sf.frb.org

Glenn D. Rudebusch
Federal Reserve Bank of San Francisco
glenn.rudebusch@sf.frb.org

This version: April 19, 2018.

Contents

This online appendix document contains all appendices referenced in the paper.

A Characteristics of Canadian Government Bonds

Tables 1 to 3 provide the contractual characteristics of all 105 Canadian government bonds in the sample and the number of monthly observations for each bond. Our sample includes all fixed-coupon bonds issued between January 2000 and April 2016 as well as 13 earlier bonds issued from 1990 to 1999. Generally, since 2008, each bond has had three or more auctions and its final total amount outstanding has been CAD 9 billion or more, which represents a substantial amount of notional.

Bond: coupon, maturity	No. obs.	Issuance		Number of auctions	Total notional amount
		Date	Amount		
(1) 10.5% 3/15/2021^{+}	196	12/15/1990	n.a.	n.a.	567
(2) 9.75% 6/1/2021^{+}	196	5/9/1991	n.a.	n.a.	286
(3) 9.25% 6/1/2022^{+}	196	12/15/1991	n.a.	n.a.	206
(4) 8% 6/1/2023^{+}	196	8/17/1992	n.a.	n.a.	2,358
(5) 9% 6/1/2025^{+}	196	8/2/1994	n.a.	n.a.	2,303
(6) 8% 6/1/2027^{+}	196	5/1/1996	n.a.	n.a.	4,036
(7) 5.75% 6/1/2029^{+}	196	2/2/1998	n.a.	n.a.	10,950
(8) 4.5% 6/1/2001	14	12/15/1998	3,500	2	7,000
(9) 5% 9/1/2004*	53	3/1/1999	2,500	4	10,850
(10) 5.25% 12/1/2001	20	6/15/1999	3,500	2	7,000
(11) 5.5% 6/1/2010†	122	8/3/1999	2,600	4	10,400
(12) 6% 9/1/2005*	65	11/15/1999	2,800	4	11,100
(13) 5.75% 6/1/2002	26	12/1/1999	3,600	2	7,200
(14) 6% 6/1/2011†	131	5/1/2000	2,600	6	15,000
(15) 6% 12/1/2002	27	6/15/2000	3,600	2	7,100
(16) 5.75% 9/1/2006*	67	11/14/2000	2,500	4	10,000
(17) 5.75% 6/1/2003	28	11/24/2000	3,500	2	7,000
(18) 5% 12/1/2003	27	6/15/2001	3,500	2	7,000
(19) 5.75% 6/1/2033^{+}	175	10/15/2001	2,000	6	10,700
(20) 5.25% 6/1/2012†	125	10/29/2001	2,500	4	9,900
(21) 4.5% 9/1/2007*	67	11/19/2001	2,500	4	9,800
(22) 3.5% 6/1/2004	28	11/30/2001	3,500	2	7,000
(23) 4.25% 12/1/2004	28	5/31/2002	3,500	2	6,500
(24) 5.25% 6/1/2013†	125	11/4/2002	2,400	4	9,600
(25) 4.25% 9/1/2008*	67	11/18/2002	2,400	4	9,400
(26) 3.5% 6/1/2005	28	11/29/2002	3,500	2	7,000
(27) 3% 12/1/2005	27	6/13/2003	2,700	2	5,200
(28) 5% 6/1/2014†	125	10/20/2003	2,400	4	9,100
(29) 4.25% 9/1/2009	67	12/1/2003	2,300	4	8,800
(30) 3% 6/1/2006	27	12/19/2003	3,500	2	7,000
(31) 3.25% 12/1/2006	28	5/28/2004	2,800	2	5,600
(32) 5% 6/1/2037^{+}	142	7/19/2004	1,500	8	11,000
(33) 4.5% 6/1/2015†	125	10/18/2004	2,100	4	8,400
(34) 4% 9/1/2010*	67	11/22/2004	2,100	4	8,100
(35) 3% 6/1/2007	27	12/10/2004	2,800	2	5,300

Table 1: **Sample of Canadian Government Fixed-Coupon Bonds**

The table reports the characteristics, first issuance date and amount, the total number of auctions, and total amount issued in millions of Canadian dollars for the Canadian government fixed-coupon bonds in the sample. Also reported are the number of monthly observation dates for each bond during the sample period from January 31, 2000, to April 29, 2016. Asterisk * indicates five-year bonds, dagger † indicates ten-year bonds, plus + indicates thirty-year bonds, and cross × indicates fifty-year bonds based on the official maturity grouping used by the Bank of Canada.

Bond: coupon, maturity	No. obs.	Issuance		Number of auctions	Total notional amount
		Date	Amount		
(36) 2.75% 12/1/2007	27	6/10/2005	3,400	2	6,800
(37) 4% 6/1/2016†	124	11/7/2005	2,100	4	8,700
(38) 3.75% 9/1/2011*	67	11/21/2005	1,900	4	8,000
(39) 3.75% 6/1/2008	27	12/16/2005	2,400	1	2,400
(40) 4.25% 12/1/2008	27	6/16/2006	3,400	2	6,800
(41) 4% 6/1/2017†	115	10/16/2006	2,300	4	9,800
(42) 3.75% 6/1/2012*	64	11/6/2006	2,000	3	6,000
(43) 3.75% 6/1/2009	28	11/17/2006	2,400	1	2,400
(44) 4.25% 12/1/2009	28	5/25/2007	3,500	2	7,100
(45) 4.25% 6/1/2018†	103	10/29/2007	2,500	5	10,100
(46) 3.75% 6/1/2010	28	11/30/2007	3,300	1	3,300
(47) 3.5% 6/1/2013*	61	2/25/2008	2,000	5	15,000
(48) 2.75% 12/1/2010	28	5/23/2008	3,500	3	11,600
(49) 4% 6/1/2041+	95	6/9/2008	1,400	10	14,100
(50) 3.75% 6/1/2019†	91	10/6/2008	2,500	5	16,000
(51) 3% 6/1/2014*	65	10/21/2008	3,000	4	16,000
(52) 1.25% 6/1/2011	26	1/27/2009	4,500	3	11,000
(53) 2% 12/1/2014*	65	4/20/2009	3,000	5	15,000
(54) 1% 9/1/2011	25	5/8/2009	3,500	3	10,000
(55) 2% 9/1/2012	37	6/1/2009	3,500	5	16,500
(56) 1.25% 12/1/2011	25	8/21/2009	3,000	3	9,500
(57) 3.5% 6/1/2020†	80	9/8/2009	3,000	4	12,500
(58) 1.5% 3/1/2012	25	11/16/2009	3,000	3	9,000
(59) 2.5% 6/1/2015*	64	11/23/2009	3,000	3	9,000
(60) 1.75% 3/1/2013	35	12/14/2009	3,200	6	18,600
(61) 1.5% 6/1/2012	24	3/12/2010	3,000	3	9,000
(62) 3% 12/1/2015*	65	4/19/2010	3,500	3	10,500
(63) 2.5% 9/1/2013	37	5/17/2010	3,000	3	9,200
(64) 3.25% 6/1/2021†	70	7/19/2010	3,000	4	11,500
(65) 1.5% 12/1/2012	25	8/13/2010	3,000	4	12,000
(66) 2% 6/1/2016*	64	11/8/2010	3,500	3	9,900
(67) 2% 3/1/2014	36	12/13/2010	3,200	3	9,600
(68) 2% 8/1/2013	25	4/8/2011	3,500	3	10,500
(69) 2.75% 9/1/2016*	61	4/26/2011	3,500	3	10,500
(70) 2.25% 8/1/2014	37	5/2/2011	3,000	5	15,600

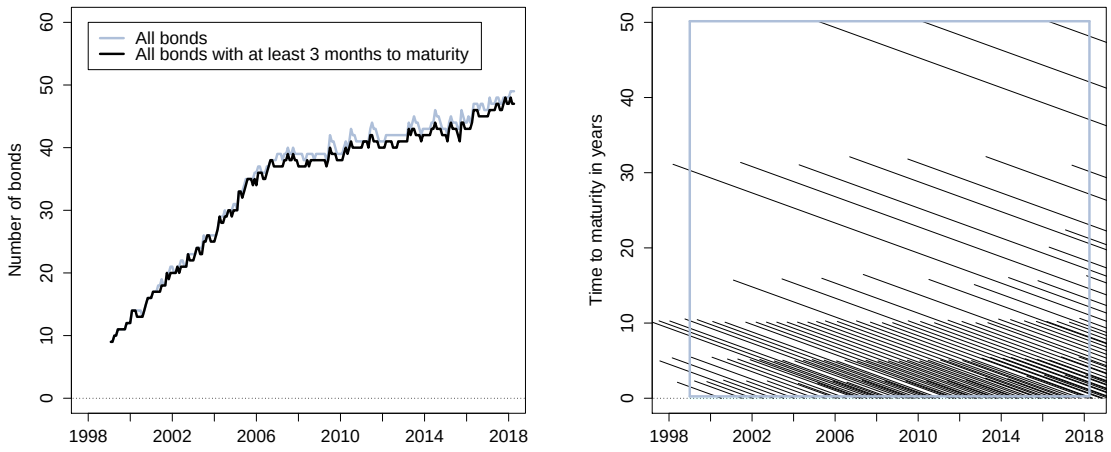
Table 2: **Sample of Canadian Government Fixed-Coupon Bonds cont.**

The table reports the characteristics, first issuance date and amount, the total number of auctions, and total amount issued in millions of Canadian dollars for the Canadian government fixed-coupon bonds in the sample. Also reported are the number of monthly observation dates for each bond during the sample period from January 31, 2000, to April 29, 2016. Asterisk * indicates five-year bonds, dagger † indicates ten-year bonds, plus + indicates thirty-year bonds, and cross × indicates fifty-year bonds based on the official maturity grouping used by the Bank of Canada.

Bond: coupon, maturity	No. obs.	Issuance		Number of auctions	Total notional amount
		Date	Amount		
(71) 3.5% 12/1/2045^{+}	59	6/13/2011	1,400	10	14,400
(72) 1.5% 11/1/2013	25	7/15/2011	3,500	3	10,500
(73) 2.75% 6/1/2022†	58	8/2/2011	2,500	5	12,700
(74) 1.5% 3/1/2017*	55	10/17/2011	3,500	3	10,500
(75) 1% 2/1/2014	25	10/21/2011	3,500	3	10,500
(76) 1% 2/1/2015	36	11/7/2011	3,000	5	15,600
(77) 0.75% 5/1/2014	24	1/13/2012	3,500	3	10,500
(78) 1.5% 8/1/2015	37	4/30/2012	2,900	5	15,300
(79) 1.5% 8/1/2017*	48	5/14/2012	3,400	3	10,200
(80) 1% 11/1/2014	26	6/22/2012	3,300	3	9,900
(81) 1.5% 6/1/2023†	46	7/30/2012	2,600	5	14,200
(82) 1.25% 2/1/2016	37	10/15/2012	2,700	5	14,700
(83) 1.25% 3/1/2018*	42	11/13/2012	3,400	3	10,200
(84) 1% 5/1/2015	24	1/18/2013	3,300	3	9,900
(85) 1% 8/1/2016	37	4/15/2013	2,700	6	17,100
(86) 1.25% 9/1/2018*	36	5/13/2013	3,400	3	10,200
(87) 2.5% 6/1/2024†	35	7/2/2013	2,800	5	13,800
(88) 1% 11/1/2015	25	7/26/2013	3,300	3	9,900
(89) 1.5% 2/1/2017	31	10/15/2013	2,700	6	17,100
(90) 1.75% 3/1/2019*	30	11/12/2013	3,400	3	10,200
(91) 1% 5/1/2016	25	1/31/2014	3,300	3	10,000
(92) 1.75% 9/1/2019*	25	4/14/2014	3,400	3	10,200
(93) 2.75% 12/1/2064$^{\times}$	25	4/28/2014	n.a.	n.a.	3,500
(94) 1.25% 8/1/2017	24	5/20/2014	2,700	6	19,100
(95) 2.75% 12/1/2048^{+}	24	6/2/2014	1,400	5	7,000
(96) 2.25% 6/1/2025†	23	6/30/2014	2,700	5	13,100
(97) 1% 11/1/2016	21	8/15/2014	3,400	3	10,200
(98) 1.5% 3/1/2020*	19	10/14/2014	3,400	3	10,200
(99) 1.25% 2/1/2018	18	11/10/2014	2,700	6	19,200
(100) 0.25% 5/1/2017	15	2/13/2015	3,400	3	10,400
(101) 0.75% 9/1/2020	13	4/13/2015	3,300	4	13,000
(102) 1.5% 6/1/2026†	10	7/21/2015	2,500	4	10,500
(103) 0.25% 11/1/2017	9	8/7/2015	3,300	4	13,400
(104) 0.75% 3/1/2021*	7	10/19/2015	3,300	4	13,800
(105) 0.25% 5/1/2018	3	2/5/2016	3,700	3	11,100

Table 3: **Sample of Canadian Government Fixed-Coupon Bonds cont.**

The table reports the characteristics, first issuance date and amount, the total number of auctions, and total amount issued in millions of Canadian dollars for the Canadian government fixed-coupon bonds in the sample. Also reported are the number of monthly observation dates for each bond during the sample period from January 31, 2000, to April 29, 2016. Asterisk * indicates five-year bonds, dagger † indicates ten-year bonds, plus + indicates thirty-year bonds, and cross × indicates fifty-year bonds based on the official maturity grouping used by the Bank of Canada.



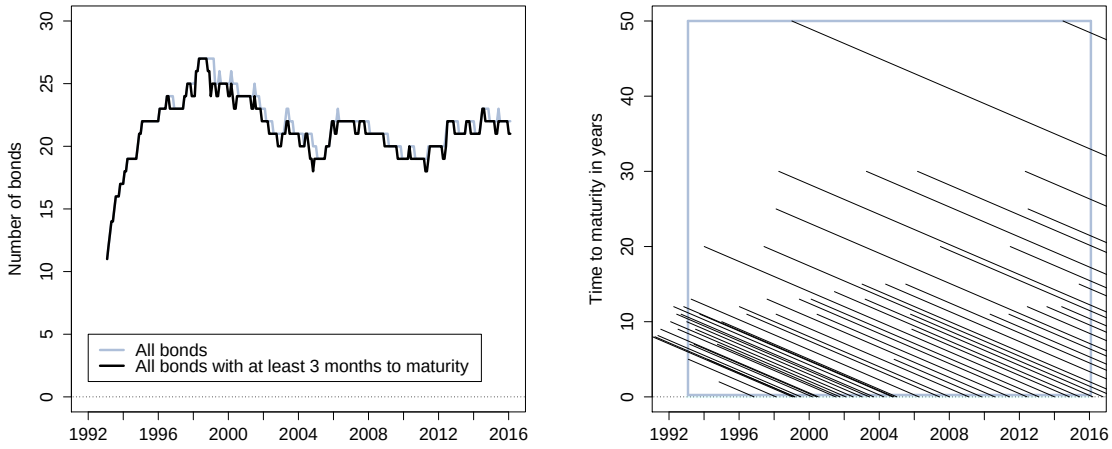
(a) Number of French bonds

(b) Maturity distribution of French bonds

Figure 1: **Description of the French Government Bond Market**
Panel (a) shows the number of French government bonds at each date. The solid grey line refers to the entire sample of bonds. The solid black line indicates the number of securities when eliminating bonds with less than three months to maturity. Panel (b) shows the maturity distribution of the full set of French government bonds. The grey rectangle indicates the considered subsample that starts in January 1999 when the euro was launched until February 2018.

A.1 Comparison to Other Sovereign Bond Markets

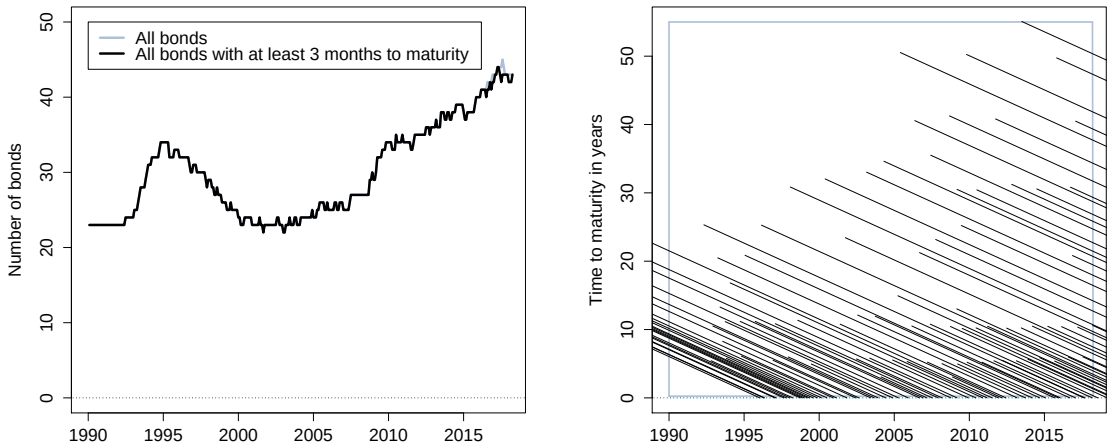
Figures 1 to 3 show the available universe of standard fixed-coupon government bonds from France, Switzerland, and the U.K. since 1999, 1993, and 1990, respectively. We note that the total number of bonds available for each country during the highlighted period in each panel is 116, 54, and 102, respectively. Thus, our Canadian sample with 105 bonds of varying maturities is very representative of the major sovereign bond markets outside the U.S.—both in terms of the number of bonds and their cross sectional distribution across time.



(a) Number of Swiss bonds

(b) Maturity distribution of Swiss bonds

Figure 2: **Description of the Swiss Government Bond Market**
Panel (a) shows the number of Swiss government bonds at each date. The solid grey line refers to the entire sample of bonds. The solid black line indicates the number of securities when eliminating bonds with less than three months to maturity. Panel (b) shows the maturity distribution of the full set of Swiss government bonds. The grey rectangle indicates the considered subsample that runs from January 1993 to January 2016.



(a) Number of U.K. gilts

(b) Maturity distribution of U.K. gilts

Figure 3: **Description of the U.K. Government Bond Market**
Panel (a) shows the number of U.K. government bonds (gilts) at each date. The solid grey line refers to the entire sample of bonds. The solid black line indicates the number of securities when eliminating bonds with less than three months to maturity. Panel (b) shows the maturity distribution of the full set of U.K. gilts. The grey rectangle indicates the considered subsample that runs from January 1990 to February 2018.

B Construction of Synthetic Zero-Coupon Yields

In this appendix, we detail the construction of the synthetic zero-coupon yield curves we consider in the paper.

B.1 The Svensson (1995) Yield Curve

This section describes how we construct synthetic zero-coupon yields using the Svensson (1995) discount function in combination with the panel of Canadian government fixed-coupon bond prices described in Appendix A. The Svensson (1995) yield curve has a flexible functional form given by

$$y_t(\tau) = \beta_0(t) + \frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau}\beta_1(t) + \left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau}\right)\beta_2(t) + \left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau}\right)\beta_3(t), \quad (1)$$

where we impose the restrictions that $\lambda_1 > \lambda_2 > 0$. This function contains the level, slope, and curvature components known from Nelson and Siegel (1987) and augments them with an additional curvature factor to provide a better fit to the long end of the yield curve. The corresponding discount function is easily obtained as $P_t^{zc}(\tau) = e^{-y_t(\tau)\tau}$. Now, consider the value at time t of a fixed-coupon bond with maturity at $t+\tau$ that pays an annual coupon C semi-annually. Its clean price, denoted $P_t(\tau, C)$, is simply the sum of its remaining cash flow payments weighted by the zero-coupon bond price function $P_t^{zc}(\tau)$:

$$P_t(\tau, C) = \frac{C}{2}\frac{(t_1 - t)}{1/2}P_t^{zc}(t_1) + \sum_{j=2}^{N}\frac{C}{2}P_t^{zc}(t_j) + P_t^{zc}(\tau), \quad t < t_1 < \ldots < t_N = \tau. \quad (2)$$

Now, the parameters in the Svensson (1995) curve, $\psi = (\beta_0, \beta_1, \beta_2, \beta_3, \lambda_1, \lambda_2)$, are estimated for each observation date by optimizing the following objective function

$$\min_{\psi}\sum_{i=1}^{n_t}\frac{1}{D_t^{Data,i}}(P_t^{Data,i} - \widehat{P_t^i}(\psi))^2, \quad (3)$$

where n_t is the number of coupon bond prices observed on day t, $P_t^{Data,i}$ is the observed price for bond number i, $\widehat{P_t^i}$ is its price implied by the Svensson (1995) discount function, and $D_t^{Data,i}$ is its duration, which is model-free and calculated before estimation based on the Macaulay formula. The stated objective is to minimize the weighted sum of the squared deviations between the actual bond prices and the predicted prices, where the weights are the

inverse of the durations of each individual security. This is identical to the objective function used by Gürkaynak et al. (2007, 2010). The optimization for each observation date is started at the same parameter vector:

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0.04173257 \\ -0.02703468 \\ -0.05262533 \\ 0.02954742 \\ 0.8378759 \\ 0.09652915 \end{pmatrix}.$$

Tables 4 and 5 report the summary statistics of the mean errors and the mean absolute errors for the ten constant-maturity zero-coupon yields constructed using the Svensson (1995) yield curve as described above when the underlying bond prices are simulated from the AFNS model as described in Section 4.1 in the paper with measurement error having a standard deviation equal to $\sigma_\varepsilon = 1$ basis point and $\sigma_\varepsilon = 10$ basis points, respectively.

The very low mean errors and mean absolute errors of the constructed zero-coupon yields for the entire cross-section of maturities across practically all $N = 100$ simulations show that the Svensson (1995) discount function is very flexible and able to produce a tight fit in a wide variety of yield environments.

Maturity	Mean errors, $\sigma_\varepsilon = 1$ basis point						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	-4.09	2.83	-9.48	-5.70	-3.75	-2.16	0.09
6	-2.37	1.69	-5.78	-3.34	-1.99	-1.21	-0.18
12	-0.32	0.42	-1.19	-0.61	-0.26	0.01	0.20
24	0.78	0.60	-0.17	0.39	0.69	1.07	1.82
36	0.43	0.47	-0.10	0.07	0.30	0.73	1.32
60	-0.48	0.37	-1.11	-0.68	-0.45	-0.25	0.05
84	-0.67	0.62	-1.94	-1.00	-0.51	-0.21	0.07
120	-0.34	0.48	-1.06	-0.71	-0.27	0.05	0.31
240	0.64	0.68	-0.36	0.07	0.57	1.02	1.91
360	-1.32	1.14	-3.12	-2.23	-1.36	-0.49	0.41

Maturity	Mean absolute errors, $\sigma_\varepsilon = 1$ basis point						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	6.79	2.55	4.00	4.80	6.15	7.84	11.57
6	3.89	1.63	2.19	2.64	3.42	4.65	6.99
12	1.05	0.40	0.57	0.76	0.95	1.28	1.83
24	1.42	0.51	0.87	1.03	1.27	1.64	2.35
36	0.99	0.45	0.49	0.66	0.84	1.23	1.84
60	1.05	0.25	0.74	0.87	1.01	1.14	1.49
84	1.26	0.56	0.62	0.81	1.08	1.56	2.34
120	0.94	0.28	0.61	0.73	0.86	1.11	1.43
240	1.46	0.53	0.85	1.04	1.38	1.72	2.64
360	3.09	0.81	1.93	2.40	2.97	3.62	4.62

Table 4: **Summary Statistics of Errors of Constructed Yields**

The top panel reports the summary statistics of the mean errors between the true zero-coupon bond yields and the zero-coupon bond yields constructed using the Svensson (1995) yield curve based on $N = 100$ simulated data sets generated by the independent-factor AFNS model, each with a distribution of coupon bond prices identical to the sample of Canadian fixed-coupon bond prices described in Section 2 of the paper and with each simulated bond price being added an *i.i.d.* measurement error with zero mean and a uniform measurement error standard deviation of $\sigma_\varepsilon = 1$ basis point scaled by the bond's duration. The bottom panel reports the corresponding summary statistics of the mean absolute errors. All numbers are measured in basis points.

Maturity	Mean errors, $\sigma_\varepsilon = 10$ basis points						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	-3.00	2.59	-7.12	-4.61	-2.79	-1.02	1.44
6	-1.74	1.54	-4.37	-2.72	-1.66	-0.58	0.71
12	-0.24	0.48	-1.07	-0.54	-0.30	0.06	0.67
24	0.52	0.57	-0.52	0.22	0.46	0.85	1.43
36	0.24	0.44	-0.37	-0.11	0.23	0.56	1.04
60	-0.34	0.45	-1.16	-0.63	-0.26	-0.03	0.25
84	-0.37	0.58	-1.56	-0.64	-0.28	0.06	0.34
120	-0.14	0.43	-0.85	-0.52	-0.04	0.18	0.43
240	0.27	0.65	-0.90	-0.18	0.31	0.61	1.19
360	-0.90	1.31	-3.07	-1.78	-1.16	0.13	1.16

Maturity	Mean absolute errors, $\sigma_\varepsilon = 10$ basis points						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	12.48	1.44	10.47	11.50	12.21	13.15	15.58
6	7.84	0.91	6.67	7.24	7.62	8.33	9.34
12	3.58	0.27	3.17	3.41	3.54	3.77	4.05
24	3.52	0.26	3.09	3.36	3.52	3.69	3.93
36	3.12	0.22	2.78	2.96	3.12	3.30	3.46
60	3.13	0.22	2.80	2.97	3.14	3.28	3.50
84	3.30	0.21	2.96	3.17	3.30	3.42	3.64
120	2.88	0.21	2.48	2.75	2.85	3.00	3.19
240	3.76	0.34	3.23	3.49	3.80	3.99	4.26
360	9.34	1.28	7.58	8.48	9.12	10.13	11.53

Table 5: **Summary Statistics of Errors of Constructed Yields**

The top panel reports the summary statistics of the mean errors between the true zero-coupon bond yields and the zero-coupon bond yields constructed using the Svensson (1995) yield curve based on $N = 100$ simulated data sets generated by the independent-factor AFNS model, each with a distribution of coupon bond prices identical to the sample of Canadian fixed-coupon bond prices described in Section 2 of the paper and with each simulated bond price being added an *i.i.d.* measurement error with zero mean and a uniform measurement error standard deviation of $\sigma_\varepsilon = 10$ basis points scaled by the bond's duration. The bottom panel reports the corresponding summary statistics of the mean absolute errors. All numbers are measured in basis points.

B.2 The Nelson and Siegel (1987) Yield Curve

This section describes how we construct synthetic zero-coupon yields using the Nelson and Siegel (1987) discount function in combination with the panel of Canadian government fixed-coupon bond prices described in Appendix A. The Nelson and Siegel (1987) yield curve has the simple, yet flexible functional form given by

$$y_t(\tau) = \beta_0(t) + \frac{1 - e^{-\lambda\tau}}{\lambda\tau}\beta_1(t) + \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)\beta_2(t), \tag{4}$$

where we impose the restriction that $\lambda > 0$.

Now, we use the same bond price equation (2) and objective function (3) as we employed above for the Svensson (1995) model to find the best fitting parameters for each set of observed bond prices. In this case, the optimization for each observation date is started at the following parameter vector:

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \lambda \end{pmatrix} = \begin{pmatrix} 0.04173257 \\ -0.02703468 \\ -0.05262533 \\ 0.4378759 \end{pmatrix}.$$

Tables 6 and 7 report the summary statistics of the mean errors and the mean absolute errors for the ten constant-maturity zero-coupon yields constructed using the Nelson and Siegel (1987) yield curve when the underlying bond prices are simulated from the AFNS model as described in Section 4.1 of the paper with measurement error having a standard deviation equal to $\sigma_\varepsilon = 1$ basis point and $\sigma_\varepsilon = 10$ basis points, respectively.

The very low mean errors and mean absolute errors of the constructed zero-coupon yields for the entire cross-section of maturities across practically all $N = 100$ simulations show that the Nelson and Siegel (1987) discount function is relatively flexible and able to produce a fit that is almost as tight as that obtained with the more flexible Svensson (1995) yield function.

Maturity	Mean errors, $\sigma_\varepsilon = 1$ basis point						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	-2.63	4.11	-10.25	-5.14	-2.31	0.08	3.66
6	-0.84	2.73	-5.86	-2.36	-0.81	1.05	3.37
12	1.05	0.92	-0.52	0.45	1.07	1.62	2.34
24	1.13	0.75	-0.10	0.53	1.04	1.73	2.35
36	-0.45	0.82	-1.69	-1.11	-0.47	-0.01	1.08
60	-2.72	0.42	-3.46	-3.05	-2.72	-2.43	-1.93
84	-2.42	0.66	-3.66	-2.92	-2.33	-1.95	-1.47
120	0.39	0.70	-0.81	-0.18	0.35	0.95	1.48
240	3.38	0.72	2.23	2.90	3.34	3.82	4.63
360	-10.06	1.75	-13.01	-11.58	-9.99	-8.70	-7.62

Maturity	Mean absolute errors, $\sigma_\varepsilon = 1$ basis point						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	11.43	1.85	8.42	9.81	11.67	12.69	14.26
6	7.60	1.15	5.71	6.72	7.76	8.37	9.34
12	2.64	0.64	1.75	2.16	2.61	3.04	3.70
24	1.97	0.42	1.38	1.60	1.92	2.28	2.71
36	2.23	0.32	1.78	1.99	2.23	2.45	2.75
60	3.02	0.38	2.42	2.76	3.01	3.22	3.70
84	2.50	0.65	1.60	2.02	2.41	2.94	3.67
120	1.68	0.28	1.28	1.45	1.66	1.87	2.10
240	3.46	0.69	2.37	2.96	3.42	3.90	4.63
360	10.58	1.51	8.44	9.41	10.37	11.76	13.09

Table 6: **Summary Statistics of Errors of Constructed Yields**

The top panel reports the summary statistics of the mean errors between the true zero-coupon bond yields and the zero-coupon bond yields constructed using the Nelson and Siegel (1987) yield curve based on $N = 100$ simulated data sets generated by the independent-factor AFNS model, each with a distribution of coupon bond prices identical to the sample of Canadian fixed-coupon bond prices described in Section 2 of the paper and with each simulated bond price being added an *i.i.d.* measurement error with zero mean and a uniform measurement error standard deviation of $\sigma_\varepsilon = 1$ basis point scaled by the bond's duration. The bottom panel reports the corresponding summary statistics of the mean absolute errors. All numbers are measured in basis points.

Maturity	Mean errors, $\sigma_\varepsilon = 10$ basis points						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	-3.39	3.85	-10.24	-5.96	-3.13	-0.89	2.35
6	-1.32	2.57	-5.86	-2.98	-1.17	0.25	2.71
12	0.90	0.92	-0.63	0.31	1.03	1.54	2.54
24	1.18	0.70	0.06	0.67	1.11	1.66	2.50
36	-0.38	0.76	-1.62	-0.92	-0.36	0.08	0.95
60	-2.67	0.46	-3.49	-2.99	-2.63	-2.33	-1.93
84	-2.36	0.67	-3.60	-2.82	-2.23	-1.96	-1.42
120	0.39	0.67	-0.76	-0.07	0.40	0.86	1.50
240	3.25	0.72	2.12	2.75	3.18	3.70	4.51
360	-9.85	1.75	-12.73	-11.27	-9.72	-8.72	-7.38

Maturity	Mean absolute errors, $\sigma_\varepsilon = 10$ basis points						
in months	Mean	Std. dev.	5 percentile	1^{st} quartile	Median	3^{rd} quartile	95 percentile
3	13.63	1.33	11.52	12.71	13.49	14.59	15.84
6	9.40	0.87	7.97	8.83	9.38	9.73	10.80
12	4.48	0.58	3.68	4.04	4.50	4.81	5.37
24	3.24	0.26	2.90	3.06	3.20	3.36	3.76
36	3.45	0.28	2.98	3.25	3.44	3.64	3.90
60	3.86	0.29	3.42	3.64	3.84	4.03	4.40
84	3.50	0.36	2.97	3.24	3.45	3.70	4.18
120	2.78	0.24	2.36	2.64	2.81	2.92	3.21
240	4.48	0.40	3.91	4.18	4.45	4.74	5.12
360	11.67	1.38	9.50	10.43	11.70	12.58	13.78

Table 7: **Summary Statistics of Errors of Constructed Yields**

The top panel reports the summary statistics of the mean errors between the true zero-coupon bond yields and the zero-coupon bond yields constructed using the Nelson and Siegel (1987) yield curve based on $N = 100$ simulated data sets generated by the independent-factor AFNS model, each with a distribution of coupon bond prices identical to the sample of Canadian fixed-coupon bond prices described in Section 2 of the paper and with each simulated bond price being added an *i.i.d.* measurement error with zero mean and a uniform measurement error standard deviation of $\sigma_\varepsilon = 10$ basis points scaled by the bond's duration. The bottom panel reports the corresponding summary statistics of the mean absolute errors. All numbers are measured in basis points.

C The Extended Kalman Filter Estimation

This appendix describes the model estimations based on the Kalman filter and the extended Kalman filter. For affine Gaussian models in general, the conditional mean vector and the conditional covariance matrix are[1]

$$\begin{aligned} E^{\mathbb{P}}[X_{t+\Delta t}|\mathcal{F}_t] &= (I - \exp(-\mathcal{K}^{\mathbb{P}}\Delta t))\theta^{\mathbb{P}} + \exp(-\mathcal{K}^{\mathbb{P}}\Delta t)X_t, \\ V^{\mathbb{P}}[X_{t+\Delta t}|\mathcal{F}_t] &= \int_t^{t+\Delta t} e^{-\mathcal{K}^{\mathbb{P}}s}\Sigma\Sigma'^{-(\mathcal{K}^{\mathbb{P}})'s}ds, \end{aligned}$$

where Δt is the time between observations. Conditional moments of discrete observations are computed and the state transition equation is obtained as

$$X_t = (I - \exp(-\mathcal{K}^{\mathbb{P}}\Delta t))\theta^{\mathbb{P}} + \exp(-\mathcal{K}^{\mathbb{P}}\Delta t)X_{t-\Delta t} + \xi_t,$$

where ξ_t refers to the Gaussian state innovations.

In the standard Kalman filter, the measurement equation is linear

$$y_t = A + BX_t + \varepsilon_t,$$

and the assumed error structure is

$$\begin{pmatrix} \xi_t \\ \varepsilon_t \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H_t \end{pmatrix}\right],$$

where the matrix H_t is assumed to be diagonal, while the matrix Q has the following structure

$$Q = \int_0^{\Delta t} e^{-\mathcal{K}^{\mathbb{P}}s}\Sigma\Sigma'^{-(\mathcal{K}^{\mathbb{P}})'s}ds.$$

In addition, the transition and measurement errors are assumed to be orthogonal to the initial states. Due to the assumed stationarity, the Kalman filter is initialized at the unconditional mean and variance of the state variables under the $\mathbb{P}$-measure, i.e., $X_0 = \theta^{\mathbb{P}}$ and $\Sigma_0 = \int_0^{\infty} e^{-\mathcal{K}^{\mathbb{P}}s}\Sigma\Sigma'^{-(\mathcal{K}^{\mathbb{P}})'s}ds$. Denote the information available at time t by $Y_t = (y_1, y_2, \ldots, y_t)$, and denote model parameters by ψ. Let $\Delta t = 1$ and consider period $t-1$ and suppose

[1] Throughout, conditional and unconditional covariance matrices are calculated using the analytical solutions provided in Fisher and Gilles (1996).

that the state update X_{t-1} and its mean square error matrix Σ_{t-1} have been obtained. The prediction step is

$$X_{t|t-1} = E^{\mathbb{P}}[X_t|Y_{t-1}] = \Phi_t^{X,0}(\psi) + \Phi_t^{X,1}(\psi)X_{t-1},$$

$$\Sigma_{t|t-1} = \Phi_t^{X,1}(\psi)\Sigma_{t-1}\Phi_t^{X,1}(\psi)' + Q_t(\psi),$$

where $\Phi_t^{X,0} = (I - \exp(-\mathcal{K}^{\mathbb{P}}\Delta t))\theta^{\mathbb{P}}$, $\Phi_t^{X,1} = \exp(-\mathcal{K}^{\mathbb{P}}\Delta t)$, and $Q_t = \int_0^{\Delta t} e^{-\mathcal{K}^{\mathbb{P}}s}\Sigma\Sigma'^{-(\mathcal{K}^{\mathbb{P}})'s}ds$.

In the time-t update step, $X_{t|t-1}$ is improved by using the additional information contained in Y_t, i.e.,

$$X_t = E[X_t|Y_t] = X_{t|t-1} + \Sigma_{t|t-1}B(\psi)'F_t^{-1}v_t,$$

$$\Sigma_t = \Sigma_{t|t-1} - \Sigma_{t|t-1}B(\psi)'F_t^{-1}B(\psi)\Sigma_{t|t-1},$$

where

$$v_t = y_t - E[y_t|Y_{t-1}] = y_t - A(\psi) - B(\psi)X_{t|t-1},$$

$$F_t = cov(v_t) = B(\psi)\Sigma_{t|t-1}B(\psi)' + H_t(\psi),$$

$$H_t(\psi) = diag(\sigma_\varepsilon^2(\tau_1), \ldots, \sigma_\varepsilon^2(\tau_{n_t})),$$

where n_t is the number of observed yields at time t.

At this point, the Kalman filter has delivered all ingredients needed to evaluate the Gaussian log likelihood, which reads

$$L^{KF}(\psi) \equiv \log l(y_1, \ldots, y_T; \psi) = \sum_{t=1}^{T} \Big(-\frac{n_t}{2}\log(2\pi) - \frac{1}{2}\log|F_t| - \frac{1}{2}v_t'F_t^{-1}v_t \Big).$$

Now, the likelihood is numerically maximized with respect to ψ using the Nelder-Mead simplex algorithm. Upon convergence, the standard errors are obtained from the estimated covariance matrix,

$$\widehat{\Omega}(\widehat{\psi}) = \frac{1}{T}\Big[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial \log l_t(\widehat{\psi})}{\partial\psi}\frac{\partial \log l_t(\widehat{\psi})}{\partial\psi}'\Big]^{-1},$$

where $\widehat{\psi}$ denotes the estimated model parameters.

In model estimations with coupon bond prices in the one-step approach and for the B-AFNS model using the two-step approach, the extended Kalman filter is needed because the measurement equations are no longer affine functions of the states. Instead, in the case of

our preferred one-step approach, the measurement equation takes the general form

$$P_t^{i,Data}(\tau, C) = \tilde{g}^i(X_t; \tau, C, \psi) + D_t^{i,Data}(\tau, C)\varepsilon_t^i,$$

which is equivalent to

$$\frac{P_t^{i,Data}(\tau, C)}{D_t^{i,Data}(\tau, C)} = \underbrace{\frac{\tilde{g}(X_t; \tau, C, \psi)}{D_t^{i,Data}(\tau, C)}}_{g^i(X_t;\tau,C,\psi)} + \varepsilon_t^i.$$

In the extended Kalman filter, this equation is linearized using a first-order Taylor expansion around the best guess of X_t in the prediction step of the Kalman filter algorithm. Thus, in the notation introduced above, this best guess is denoted $X_{t|t-1}$ and the approximation is given by

$$g^i(X_t; \tau, C, \psi) \approx g^i(X_{t|t-1}; \tau, C, \psi) + \frac{\partial g^i(X_t; \tau, C, \psi)}{\partial X_t}\Big|_{X_t = X_{t|t-1}} (X_t - X_{t|t-1}).$$

Thus, by defining

$$A_t^i(\psi) \equiv g^i(X_{t|t-1}; \tau, C, \psi) - \frac{\partial g^i(X_t; \tau, C, \psi)}{\partial X_t}\Big|_{X_t = X_{t|t-1}} X_{t|t-1} \quad \text{and} \quad B_t^i(\psi) \equiv \frac{\partial g^i(X_t; \tau, C, \psi)}{\partial X_t}\Big|_{X_t = X_{t|t-1}},$$

the measurement equation can be given an affine form as

$$\frac{P_t^{i,Data}(\tau, C)}{D_t^{i,Data}(\tau, C)} = A_t^i(\psi) + B_t^i(\psi)X_t + \varepsilon_t^i,$$

and the steps in the algorithm proceed as previously described, except that the standard errors are obtained from

$$\widehat{\Omega}^{QML}\left(\widehat{\psi}\right) = \frac{1}{T}\overline{H}(\widehat{\psi})^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial \log l_t(\widehat{\psi})}{\partial \psi}\frac{\partial \log l_t(\widehat{\psi})}{\partial \psi}'\right]\overline{H}(\widehat{\psi})^{-1},$$

where $\overline{H}(\psi)$ is the Hessian matrix evaluated as described in Harvey (1989).

D One-Step Approach with Alternative Measurement Equations

In this appendix, we analyze two alternative formulations of the measurement equation in the one-step approach. The first alternative specification of the measurement equation that we consider is to omit scaling the pricing errors by duration, as in Pancost (2018). That is, we let

$$P_t^{i,Data}(\tau, C) = P_t^i(\tau, C) + \varepsilon_t^i, \tag{5}$$

where $\varepsilon_t^i \sim \mathcal{NID}\left(0, \sigma_\varepsilon^2\right)$, although the variance of ε_t^i is likely to depend on the time to maturity of the ith bond. Our second alternative to equation (5) in the paper is to express bond prices in terms of their implied yield to maturity based on equation (6) in the paper. That is, we consider the measurement equation

$$y_t^{i,Data}(\tau, C) = y_t^i(\tau, C) + \varepsilon_t^i, \tag{6}$$

where $y_t^{i,Data}(\tau, C)$ and $y_t^i(\tau, C)$ are the implied yield to maturity on the observed and model-implied bond price, respectively, and $\varepsilon_t^i \sim \mathcal{NID}\left(0, \sigma_\varepsilon^2\right)$. The appealing feature of this second specification is that coupon bonds are represented in yields and not prices, and this makes the one-step approach more similar to the traditional two-step approach, where the measurement equation is formulated in terms of synthetic zero-coupon yields. The main drawback of equation (6) is its computational requirement, as we need to solve the nonlinear fixed-point problem in equation (7) in the paper for *every* evaluation of the measurement equation in the EKF.[2]

Table 8 shows that the estimates of the AFNS model are very similar when formulating the measurement equation in terms of either duration-scaled pricing errors or the implied yield to maturity, although the persistence ($\kappa_{11}^{\mathbb{P}}$) and mean ($\theta_1^{\mathbb{P}}$) of the level factor differ somewhat. When omitting the duration-scaling of the pricing errors, the slope factor becomes more persistent and less volatile, while the curvature factor becomes much less persistent and much more volatile compared to any of the two other formulations of the measurement equation. Also, the estimate of λ is significantly lower when using equation (5), implying that the AFNS

[2] In our implementation of the AFNS model, the computing time for one evaluation of the EKF is roughly 4.5 times longer when the measurement equation is formulated in yields to maturity as opposed to duration-scaled pricing errors.

Par.	Bond prices				Bond yields to maturity	
	With scaling		Without scaling			
	Est	SE	Est	SE	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.1060	0.0763	0.0233	0.0302	0.0055	0.0044
$\kappa_{22}^{\mathbb{P}}$	0.2157	0.1443	0.1542	0.1910	0.2428	0.1363
$\kappa_{33}^{\mathbb{P}}$	0.7255	0.3649	2.1973	0.2999	0.7177	0.3786
σ_{11}	0.0052	0.0001	0.0020	0.0004	0.0051	0.0002
σ_{22}	0.0103	0.0010	0.0092	0.0006	0.0102	0.0010
σ_{33}	0.0207	0.0015	0.0490	0.0016	0.0209	0.0014
$\theta_1^{\mathbb{P}}$	0.0529	0.0034	0.0938	0.0045	0.0962	0.0474
$\theta_2^{\mathbb{P}}$	-0.0275	0.0093	-0.0636	0.0174	-0.0260	0.0050
$\theta_3^{\mathbb{P}}$	-0.0230	0.0060	-0.0508	0.0057	-0.0200	0.0051
λ	0.3747	0.0105	0.1415	0.0026	0.3722	0.0120

Table 8: **Estimates of the AFNS Model with Different Measurement Equations**
This table reports the estimated parameters (Est) in the AFNS model with independent factors in the one-step approach using QML with different formulations of the measurement equation as described in the main text. The standard errors (SE) are computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, computed as outlined in Harvey (1989). The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

model puts more weight on matching bonds with a long time to maturity. This is also evident from Table 9, where the one-step approach without duration-scaled pricing errors gives very low fitting errors for long-term bonds but also very large errors for short-term bonds. In contrast, the fit obtained when using the implied yield to maturity in equation (6) is almost identical to our benchmark specification in equation (5) in the paper.

Table 10 reports the standard deviations of the pricing errors within each maturity bucket for our benchmark estimates of the AFNS model in the one-step approach. When scaling pricing errors by duration, we find that these standard deviations are more or less constant across the maturity buckets, except possibly for the last bucket with the longest maturities. In contrast, the pricing errors increase gradually with maturity when ε_t^i is not scaled by duration and hence violates the assumption that the standard deviations of ε_t^i are constant across bonds. Another way to justify our benchmark specification in equation (5) in the paper is to briefly consider the more general measurement equation $P_t^{i,Data}(\tau, C) = P_t^i(\tau, C) + D_t^{i,Data}(\tau, C)^{\beta_\varepsilon}\varepsilon_t^i$, where $\beta_\varepsilon \in \mathbb{R}_+$ controls how the pricing errors are scaled by duration. The estimates of the AFNS model in the one-step approach with this more general measurement equation are basically identical to those reported in Table 2 in the paper, as $\hat{\beta}_\varepsilon = 1.00003$ and with a standard error of 3.85×10^{-5} for $\hat{\beta}_\varepsilon$. Hence, we cannot reject the null hypothesis

Maturity bucket in years	No. obs.	Bond prices				Bond yields to maturity	
		With scaling		Without scaling			
		Mean	RMSE	Mean	RMSE	Mean	RMSE
0-2	1,472	-0.09	5.75	-1.66	22.84	-0.13	5.57
2-4	1,098	0.46	4.76	-0.96	8.46	0.40	4.43
4-6	744	-0.32	4.01	1.22	5.56	-0.15	3.52
6-8	404	-1.24	5.46	0.34	6.17	-0.83	5.14
8-10	477	-2.54	6.07	-1.67	6.89	-1.96	5.73
10-12	289	-1.07	6.36	-2.11	7.61	-0.17	6.45
12-14	155	3.78	6.72	1.63	3.63	5.32	8.85
14-16	168	0.76	4.32	-0.24	2.80	1.49	5.70
16-18	179	0.70	4.66	-0.75	4.09	1.43	6.01
18-20	192	1.71	4.33	0.17	3.98	2.65	5.42
20-22	186	3.64	4.98	2.53	4.60	4.59	5.82
22-24	142	0.84	4.74	1.80	4.73	1.36	4.24
24-26	124	0.05	5.63	2.37	4.21	1.08	4.64
26-28	113	-5.45	8.71	-0.60	3.79	-4.65	6.82
28<	288	-4.97	11.98	-0.95	3.66	-4.29	10.42
All bonds	6,031	-0.33	5.90	-0.51	12.54	0.01	5.65

Table 9: **Summary Statistics of Bond Fitted Errors in the AFNS Model with Different Measurement Equations**

This table reports the mean pricing errors (Mean) and the root mean-squared pricing errors (RMSE) of the Canadian bond prices for the AFNS model with independent factors estimated with the one-step approach using QML and different formulations of the measurement equation as described in the main text. The pricing errors are reported in basis points and computed as the difference between the implied yield on the coupon bond and the model-implied yield on this bond. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

	Maturity buckets in years														
	0-2	2-4	4-6	6-8	8-10	10-12	12-14	14-16	16-18	18-20	20-22	22-24	24-26	26-28	28-
Duration scaling	5.83	4.89	4.18	6.06	6.65	8.06	9.09	6.80	7.36	6.18	6.99	7.59	7.48	7.90	11.90
No duration scaling	5.58	13.30	19.15	37.01	48.97	66.26	78.64	68.06	77.82	69.05	78.37	94.34	106.78	120.46	253.11

Table 10: **Standard Deviations of Pricing Errors**

This table reports the standard deviations of pricing errors for the AFNS model with independent factors estimated with the one-step approach using QML and the measurement equation with duration-scaled pricing errors. The pricing errors are evaluated at the filtered states and computed as the difference between observed and model-implied coupon bond prices, when including and omitting the duration scaling, respectively. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

that $\beta_\varepsilon = 1$ (p-value of 34%) and therefore our benchmark specification in equation (5) in the paper. Note also that we clearly reject the null hypothesis that $\beta_\varepsilon = 0$, as implied by the specification in equation (5).

E One-Step Approach with Other Nonlinear Filters

In this appendix, we provide the details of the implementation for two alternative nonlinear filtering methods, namely the unscented Kalman filter as described in Filipović and Trolle (2013) and the SR approach developed in Andreasen and Christensen (2015).

E.1 The Unscented Kalman Filter

Recall from Appendix C that the prediction step in the Kalman filter algorithm for Gaussian models like ours is

$$X_{t|t-1} = E^{\mathbb{P}}[X_t|Y_{t-1}] = \Phi_t^{X,0} + \Phi_t^{X,1} X_{t-1},$$

$$\Sigma_{t|t-1} = \Phi_t^{X,1} \Sigma_{t-1} (\Phi_t^{X,1})' + Q_t,$$

where $\Phi_t^{X,0} = (I - \exp(-\mathcal{K}^{\mathbb{P}}\Delta t))\theta^{\mathbb{P}}$, $\Phi_t^{X,1} = \exp(-\mathcal{K}^{\mathbb{P}}\Delta t)$, and $Q_t = \int_0^{\Delta t} e^{-\mathcal{K}^{\mathbb{P}}s} \Sigma\Sigma' e^{-(\mathcal{K}^{\mathbb{P}})'s} ds$.

Next, calculate the mean and covariance matrix of the prediction errors

$$v_t = y_t - E[y_t|Y_{t-1}] = y_t - A - BX_{t|t-1},$$

$$F_t = cov(v_t) = B\Sigma_{t|t-1}B' + H_t,$$

$$H_t = diag(\sigma_\varepsilon^2(\tau_1), \ldots, \sigma_\varepsilon^2(\tau_{n_t})).$$

In the time-t update step, $X_{t|t-1}$ is improved by using the additional information contained in Y_t, i.e.,

$$X_t = E[X_t|Y_t] = X_{t|t-1} + \Sigma_{t|t-1}B'F_t^{-1}v_t,$$

$$\Sigma_t = \Sigma_{t|t-1} - \Sigma_{t|t-1}B'F_t^{-1}B\Sigma_{t|t-1}.$$

In model estimations with coupon bond prices in the one-step approach and for the B-AFNS model using the two-step approach, the extended Kalman filter is needed because the measurement equations are no longer affine functions of the states. Instead, the measurement equation takes the general form

$$Z_t = g(X_t; \tau, C, \psi) + \varepsilon_t.$$

To begin the description of the unscented Kalman filter, let L be the dimension of $X_{t|t-1}$

and let κ be a scaling parameter that we fix at $\kappa = 0.5$.

Now, a set of $2L+1$ sigma points and associated weights are selected based on the scheme

$$\begin{aligned}
\widehat{X}^0_{t|t-1} &= X_{t|t-1}, \quad w^0 = \frac{\kappa}{L+\kappa}; \\
\widehat{X}^i_{t|t-1} &= X_{t|t-1} + \sqrt{(L+\kappa)\Sigma_{t|t-1}}_i, \quad w^i = \frac{1}{2(L+\kappa)}, \quad i = 1, \ldots, L; \\
\widehat{X}^i_{t|t-1} &= X_{t|t-1} - \sqrt{(L+\kappa)\Sigma_{t|t-1}}_i, \quad w^i = \frac{1}{2(L+\kappa)}, \quad i = L+1, \ldots, 2L.
\end{aligned}$$

Here, $\sqrt{(L+\kappa)\Sigma_{t|t-1}}_i$ is the ith column of the matrix square root.[3]

In the prediction step, the predicted value of the measurement now takes the form

$$\widehat{Z}_{t|t-1} = \sum_{i=0}^{2L} w^i g(\widehat{X}^i_{t|t-1}; \tau, C, \psi)$$

and the covariance matrix of the prediction error is given by

$$F_{t|t-1} = \sum_{i=0}^{2L} w^i \Big(g(\widehat{X}^i_{t|t-1}; t^i_0, \tau^i) - \widehat{Z}_{t|t-1} \Big) \Big(g(\widehat{X}^i_{t|t-1}; t^i_0, \tau^i) - \widehat{Z}_{t|t-1} \Big)' + H_t.$$

Finally, for the update step, we need

$$W_t = \sum_{i=0}^{2L} w^i \Big(\widehat{X}^i_{t|t-1} - X_{t|t-1} \Big) \Big(g(\widehat{X}^i_{t|t-1}; t^i_0, \tau^i) - \widehat{Z}_{t|t-1} \Big)' \Sigma^{-1}_{t|t-1},$$

which is followed by the actual updating

$$X_t = X_{t|t-1} + W_t(Z_t - \widehat{Z}_{t|t-1})$$

and

$$\Sigma_t = \Sigma_{t|t-1} - W_t F^{-1}_{t|t-1} W_t'.$$

Now, everything else in the Kalman filter algorithm proceeds as described in Appendix C.

[3]Since $\Sigma_{t|t-1}$ is symmetric, we have that $\sqrt{(L+\kappa)\Sigma_{t|t-1}} = VSV^T$, where S is a diagonal matrix with the square root of the eigenvalues of $(L+\kappa)\Sigma_{t|t-1}$, while V contains its eigenvectors.

E.2 The SR Approach

To describe the SR approach, it is necessary to define two vectors containing partly overlapping subsets of parameters. First, θ_1 denotes the 'risk-neutral parameters' that determine the relationship between the factors and yields, while θ_2 denotes the 'time series parameters' that determine the $\mathbb{P}$-dynamics of the states. Since Σ appears in both θ_1 and θ_2, it is necessary to further partition these vectors as $\theta_1 \equiv \left[\begin{array}{cc} \lambda & vech\left(\Sigma\right)' \end{array} \right]'$ and $\theta_2 \equiv \left[\begin{array}{cc} \theta_{22}' & vech\left(\Sigma\right)' \end{array} \right]'$ for the AFNS model. The vector θ_{22} is given by $\left[\begin{array}{cc} h_0' & vec\left(h_x\right)' \end{array} \right]'$, as we initially estimate the discrete-time VAR(1) model:

$$X_{t+1} = h_0 + h_x X_t + \Sigma^{VAR} \epsilon_{t+1}. \tag{7}$$

Below, we show how to recover the corresponding diffusion coefficients for the state process—including Σ in the diffusion.

The SR approach has three steps. In Step 1, we jointly estimate θ_1 and the states using cross-section regressions. For a given value of θ_1, we estimate the factors in period t using a cross-section regression, i.e. $\hat{X}_t\left(\theta_1\right) = \arg\min_{X_t \in \mathbb{R}^{n_x}} \frac{1}{2n_{y,t}} \sum_{i=1}^{n_{y,t}} \left(y_{t.i} - g^i(X_t; \theta_1, \tau, C)\right)^2$. To estimate θ_1, we therefore pool the squared residuals from these cross-section regressions and minimize their sum with respect to θ_1, i.e.

$$\hat{\theta}_1^{step1} = \arg\min_{\theta_1 \in \Theta_1} \frac{1}{2N} \sum_{t=1}^{T} \sum_{i=1}^{n_{y,t}} \left(y_{t.i} - g^i(\hat{X}_t\left(\theta_1\right); \theta_1, \tau, C)\right)^2. \tag{8}$$

Here, $\hat{\theta}_1^{step1}$ denotes the Step 1 estimate of θ_1 , $N \equiv \sum_{t=1}^{T} n_{y,t}$, and Θ_1 is the feasible domain of θ_1.

In Step 2 of the SR approach, we estimate θ_2 using the estimated states $\left\{\hat{X}_t\left(\hat{\theta}_1^{step1}\right)\right\}_{t=1}^{T}$. As shown in Andreasen and Christensen (2015), when θ_2 is unrestricted, it is possible to estimate (7) by running a modified regression with all second moments corrected for estimation uncertainty in the factors.

In Step 3 of the SR approach, we combine the estimates of Σ from Step 1 and 2 ($\hat{\Sigma}^{step1}$ and $\hat{\Sigma}^{step2}$, respectively) optimally and re-estimate λ conditional on the optimal estimate of Σ. However, $\hat{\Sigma}^{step1}$ is estimated very inaccurately compared to $\hat{\Sigma}^{step2}$, and we therefore follow Andreasen and Christensen (2015) and simply use the time series estimate $\hat{\Sigma}^{step2}$ for Step 3,

where we condition on $\hat{\Sigma}^{step2}$ and re-estimate λ as

$$\hat{\lambda}^{step3} = \arg \min_{\lambda \in \Theta_{11}} \frac{1}{2N} \sum_{t=1}^{T} \sum_{i=1}^{n_{y,t}} \left(y_{t.i} - g^i(\hat{X}_t \left(\lambda, \hat{\Sigma}^{step2} \right); \lambda, \hat{\Sigma}^{step2}, \tau, C) \right)^2, \tag{9}$$

where $\hat{\lambda}^{step3}$ denotes the Step 3 estimate of λ and Θ_{11} is the feasible domain of λ. We finally update our estimate of θ_2 by re-running Step 2 using the estimated states $\left\{ \hat{X}_t \left(\hat{\lambda}^{step3}, \hat{\Sigma}^{step2} \right) \right\}_{t=1}^{T}$.

To see how estimates of the discrete-time VAR(1) model are related to the diffusion parameters of interest, let us re-state the discrete-time VAR(1) model as

$$X_{t+\Delta t} = h_0 + h_x X_{\Delta t} + \Sigma^{VAR} \epsilon_{t+\Delta t}$$

for a given value of Δt. Here, $\epsilon_{t+1} \sim \mathcal{NID}(0, I)$. Now, recall that the exact distribution for the first and second moments in the diffusion process for X_t under the $\mathbb{P}$-measure are

$$\begin{aligned} E_t^{\mathbb{P}}[X_{t+\Delta t}] &= \left(I - \exp\left\{ -\mathcal{K}^{\mathbb{P}} \Delta t \right\} \right) \theta^{\mathbb{P}} + \exp\left\{ -\mathcal{K}^{\mathbb{P}} \Delta t \right\} X_t, \\ V_t^{\mathbb{P}}[X_{t+\Delta t}] &= \int_0^{\Delta t} \exp\left\{ -\mathcal{K}^{\mathbb{P}} s \right\} \Sigma \Sigma' \exp\left\{ -\mathcal{K}^{\mathbb{P}} s \right\}' ds, \end{aligned}$$

where $\theta^{\mathbb{P}}$, $\mathcal{K}^{\mathbb{P}}$, and Σ are the parameters in the diffusion. The procedure to compute $V_t^{\mathbb{P}}[X_{t+\Delta t}]$ is derived in Fisher and Gilles (1996) and is as follows:

- Diagonalize $\mathcal{K}^{\mathbb{P}}$, i.e. compute $[V, E] = eig\left(\mathcal{K}^{\mathbb{P}}\right)$, where V are the eigenvectors and E are the eigenvalues in a diagonal matrix.
- Compute $S = V^{-1} \Sigma \Sigma' \left(V^{-1}\right)'$.
- Then, $V_t^{\mathbb{P}}[X_{t+\Delta t}]_{i,j} = S_{i,j} \left(\frac{1 - \exp\{-(E_i + E_j)\Delta t\}}{E_i + E_j} \right)$, provided $E_i + E_j \neq 0$.

Hence, we have

$$h_x = \exp\left\{ -\mathcal{K}^{\mathbb{P}} \Delta t \right\},$$

or, equivalently,

$$\mathcal{K}^{\mathbb{P}} = \frac{-\log^M(h_x)}{\Delta t},$$

where $\log^M(h_x)$ denotes the matrix logarithm of h_x as given by logm in MATLAB. Note that if the VAR(1) model is estimated on monthly data, then we let $\Delta t = 1/12$.[4] To get the

[4]Note that this is not necessarily feasible for arbitrary h_x matrices, but it is valid in our case where h_x is restricted to being diagonal.

intercept, we have

$$\begin{aligned} h_0 &= \left(I - \exp\left\{-\mathcal{K}^{\mathbb{P}}\Delta t\right\}\right)\theta^{\mathbb{P}} \\ &= (I - h_x)\,\theta^{\mathbb{P}}, \end{aligned}$$

which is easily re-arranged to yield

$$\theta^{\mathbb{P}} = (I - h_x)^{-1} h_0.$$

Finally, to get Σ in the diffusion, we observe that

$$\left[V^{-1}\Sigma\Sigma'\left(V^{-1}\right)'\right]_{i,j}\left(\frac{1-\exp\left\{-\left(E_i+E_j\right)\Delta t\right\}}{E_i+E_j}\right) = \left[\Sigma^{VAR}\left(\Sigma^{VAR}\right)'\right]_{i,j}$$

or, equivalently,

$$\left[V^{-1}\Sigma\Sigma'\left(V^{-1}\right)'\right]_{i,j} = \left[\frac{E_i+E_j}{1-\exp\left\{-\left(E_i+E_j\right)\Delta t\right\}}\Sigma^{VAR}\left(\Sigma^{VAR}\right)'\right]_{i,j}$$

for all i and j.

Now, let

$$\Omega \equiv \left[\frac{E_i+E_j}{1-\exp\left\{-\left(E_i+E_j\right)\Delta t\right\}}\Sigma^{VAR}\left(\Sigma^{VAR}\right)'\right]_{i,j}.$$

Then

$$V^{-1}\Sigma\Sigma'\left(V^{-1}\right)' = \Omega \quad \Longleftrightarrow \quad \Sigma\Sigma' = V\Omega V'.$$

Hence, we can obtain Σ by $chol\left(V\Omega V', 'lower'\right)$ in MATLAB.

E.3 Results

Table 11 shows that the estimates of the AFNS model are hardly affected by replacing the EKF with the UKF. This indicates that the nonlinearities in the measurement equation from pricing coupon bonds are very small and therefore approximated very accurately by the local linearization used in the EKF. A similar finding is reported in Christensen and Rudebusch (2015). As a result, the two filters give nearly the same fit to bond prices in the AFNS model, as the overall RMSE is 5.87 and 5.90 basis points with the UKF and EKF, respectively. The EKF is computationally less demanding to implement than the UKF, particularly when using

Par.	EKF		Unscented KF		SR approach	
	Est	SE	Est	SE	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.1060	0.0763	0.1515	0.0558	0.1258	0.1322
$\kappa_{22}^{\mathbb{P}}$	0.2157	0.1443	0.2172	0.1440	0.3406	0.1819
$\kappa_{33}^{\mathbb{P}}$	0.7255	0.3649	0.7298	0.3503	0.7826	0.3719
σ_{11}	0.0052	0.0001	0.0052	0.0001	0.0060	0.0004
σ_{22}	0.0103	0.0010	0.0103	0.0010	0.0108	0.0010
σ_{33}	0.0207	0.0015	0.0208	0.0015	0.0193	0.0016
$\theta_1^{\mathbb{P}}$	0.0529	0.0034	0.0494	0.0056	0.0394	0.0181
$\theta_2^{\mathbb{P}}$	-0.0275	0.0093	-0.0282	0.0087	-0.0354	0.0086
$\theta_3^{\mathbb{P}}$	-0.0230	0.0060	-0.0236	0.0056	-0.0264	0.0071
λ	0.3747	0.0105	0.3746	0.0105	0.3329	0.0004

Table 11: **Different Nonlinear Filter Estimates of the AFNS Model**
This table reports the estimated parameters (Est) in the AFNS model with independent factors in the one-step approach, using either QML based on the EKF, QML based on the Uncented Kalman filter, or the SR approach. The standard errors (SE) for two QML estimators are computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, as outlined in Harvey (1989). The SE for the SR approach are robust to heteroskedasticity and autocorrelations in the state invations, as well as pricing errors with heteroskedasticity in the time series dimension, cross-sectional correlation (when bonds are sorted by duration), and autocorrelation which we implement with w_D= 5 and w_T= 10 in the provided estimator of Andreasen and Christensen (2015). The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

an analytical expression for the Jacobian in the EKF, and we therefore prefer using the EKF in the one-step approach.

We see somewhat larger differences when comparing the estimates from the EKF to those implied by the SR approach, but we also note that many of these differences are within the confidence intervals for the estimated parameters and hence not statistical significant. The standard errors in Table 11 also reveal that the estimates from the SR approach are generally less efficient than those provided by the EKF, which is consistent with the the simulation results in Andreasen and Christensen (2015). The main exception to this pattern is the precision of the λ estimates, where the QML standard error is 0.0105 but only 0.0004 in the SR approach, which properly corrects for pricing errors displaying heteroskedasticity in the time series dimension, cross-sectional correlation (when bonds are sorted by duration), and autocorrelation. The close similarlity between the two set of estimates is also evident from the goodness of fit, as the SR approach gives an overall RMSE of 5.87 basis points, and hence basically provides the same satisfying fit as the EKF. Thus, although the SR approach is less restrictive in terms of the admissible specifications of the measurement error distribution

Maturity bucket in years	No. obs.	Standard ext. KF		Unscented KF		SR approach	
		Mean	RMSE	Mean	RMSE	Mean	RMSE
0-2	1,472	-0.09	5.75	-0.14	5.76	-0.18	5.76
2-4	1,098	0.46	4.76	0.34	4.75	0.17	4.88
4-6	744	-0.32	4.01	-0.49	4.03	0.01	3.91
6-8	404	-1.24	5.46	-1.44	5.51	-0.64	5.41
8-10	477	-2.54	6.07	-2.76	6.17	-2.41	6.14
10-12	289	-1.07	6.36	-1.32	6.36	-1.50	6.34
12-14	155	3.78	6.72	3.54	6.58	3.06	6.34
14-16	168	0.76	4.32	0.51	4.28	-0.17	4.31
16-18	179	0.70	4.66	0.45	4.62	-0.29	4.50
18-20	192	1.71	4.33	1.45	4.24	0.71	4.14
20-22	186	3.64	4.98	3.30	5.33	2.31	4.16
22-24	142	0.84	4.74	0.45	5.06	-0.07	4.44
24-26	124	0.05	5.63	-0.29	5.29	-0.67	5.20
26-28	113	-5.45	8.71	-5.82	8.69	-5.17	7.72
28<	288	-4.97	11.98	-5.33	12.05	-3.11	12.87
All bonds	6,031	-0.33	5.90	-0.50	5.87	-0.42	5.87

Table 12: **Summary Statistics of Bond Fitted Errors in the AFNS Model with Alternative Nonlinear Filters**

This table reports the mean pricing errors (Mean) and the root mean-squared pricing errors (RMSE) of the Canadian bond prices for the AFNS model with independent factors estimated with the one-step approach using QML and three different nonlinear filters as described in the text. The pricing errors are reported in basis points and computed as the difference between the implied yield on the coupon bond and the model-implied yield on this bond. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

compared to conventional Kalman filter estimations, its convenience comes at the cost of being less efficient. However, this difference does not seem to affect our results.

F The One-Step Approach and Full ML Estimation

It is well-known that the adopted QML estimator in the one-step approach based on the EKF (and the UKF) may induce an efficiency loss compared to ML, but it is perhaps less recognized that consistency of this QML estimator cannot be established as the sample size T tends to infinity.[5] To explore whether the performance of the one-step approach can be improved by adopting a better estimator, we next show how the one-step approach can be implemented with a fully efficient ML estimator.

We have so far adopted a Bayesian perspective when filtering out the states in both the one- and two-step approach. But the one-step approach is characterized by a large set of observables in the cross-sectional dimension, and it therefore seems natural to adopt a classical perspective to filtering, as commonly considered in the estimation of large factor models (see, for instance, Bai and Ng (2002) and Bai (2003)).[6] That is, we now consider the states $X_{1:T} \equiv \{X_t'\}_{t=1}^T$ as parameters along with the model parameters ψ.[7] The main advantage of considering the states $X_{1:T}$ as parameters is that the likelihood function can be evaluated without simulation for a nonlinear DTSM with Gaussian state innovations and pricing errors, and this in turn makes full ML estimation feasible within the one-step approach.

To realize this, let $\tilde{\psi} \equiv \begin{bmatrix} \psi & X_{1:T} \end{bmatrix}$ denote the extended set of model parameters and let n_t denote the number of bond prices in period t, which we collect in Y_t. Hence, $Y_{1:T} \equiv \{Y_t\}_{t=1}^T$ refers to the entire sample of bond prices. The relation between bond prices and the states is then expressed in a condensed manner by the measurement equation

$$Y_t = g\left(X_t; \psi\right) + \varepsilon_t, \tag{10}$$

where $g\left(X_t; \psi\right)$ is a nonlinear function in X_t and $\varepsilon_t \sim \mathcal{NID}\left(0, R_{\varepsilon,t}\right)$.[8] The state transition

[5] This is because the approximated nature of the EKF (and the UKF) implies that the conditional first and second moments for the prediction errors related to coupon bond prices cannot be computed *exactly* at the true model parameters, see Bollerslev and Wooldridge (1992) and Andreasen (2013).

[6] A classical perspective to filtering is also adopted in the SR approach of Andreasen and Christensen (2015) and in Andersen et al. (2015) when estimating option pricing models.

[7] The curve-fitting procedure of Svensson (1995), Bliss (1996), and Gürkaynak et al. (2007, 2010) among others adopt the same classical perspective, as they estimate a parametric model for a daily yield curve, where the "states" in these curves are treated as parameters and estimated from a large panel of bond prices.

[8] The subscript t on $R_{\varepsilon,t}$ indicates that its dimension adapts to the available number of bonds throughout the sample.

dynamics under the $\mathbb{P}$-measure is after an appropriate Euler-discretization given by

$$X_{t+1} = h\left(X_t;\psi\right) + w_{t+1}, \tag{11}$$

where $h\left(X_t;\psi\right)$ is a potentially nonlinear function in X_t and $w_{t+1} \sim \mathcal{NID}\left(0, R_w\right)$. Given the imposed distributional assumptions on the system in equations (10) and (11), the log-likelihood function $L\left(\tilde{\psi}|Y_{1:T}\right)$ is then proportional to (see Durbin and Koopman (2001))

$$\begin{aligned} L\left(\tilde{\psi}|Y_{1:T}\right) \quad \propto \quad & \frac{T}{2}\log\left|R_w^{-1}\right| - \frac{1}{2}\sum_{t=1}^{T}\left(X_{t+1} - h\left(X_t;\psi\right)\right)' R_w^{-1}\left(X_{t+1} - h\left(X_t;\psi\right)\right) \\ & + \sum_{t=1}^{T}\frac{1}{2}\log\left|R_{\varepsilon,t}^{-1}\right| - \frac{1}{2}\sum_{t=1}^{T}\left(Y_t - g\left(X_t;\psi\right)\right)' R_{\varepsilon,t}^{-1}\left(Y_t - g\left(X_t;\psi\right)\right). \end{aligned} \tag{12}$$

The ML estimator is then given by

$$\widehat{\tilde{\psi}}_{ML} = \arg\max_{\tilde{\psi}\in\tilde{\Psi}} L\left(\tilde{\psi}|Y_{1:T}\right), \tag{13}$$

where $\tilde{\Psi}$ denotes the feasible set for $\tilde{\psi}$. To make this optimization problem computationally feasible, we use the procedure in Durbin and Koopman (2001) to numerically concentrate out $X_{1:T}$ from $L\left(\tilde{\psi}|Y_{1:T}\right)$ for a given value of ψ. As explained below, this is done by iterating the Kalman filter and smoother on a linearized version of the system in equations (10) and (11), where convergence for the AFNS model typically is achieved within five iterations.[9] The asymptotic distribution of $\hat{\psi}_{ML}$ when $n_t \longrightarrow \infty$ for all t and $T \longrightarrow \infty$ at the same rate as n_t is multivariate normal, and the standard errors are given by the inverse of the Hessian matrix or the variance of the score for the concentrated log-likelihood function (see Hahn and Newey (2004)).[10]

We next provide the details regarding the this fully efficient ML estimator for the one-step approach. To describe the procedure for concentrating out $X_{1:T}$ of the log-likelihood function

[9]The specification in (11) omits nonlinearities between the states and the innovations, but this is without loss of generality, as shown in Appendix G. Hence, the proposed ML estimator may also be applied to DTSMs with stochastic volatility.

[10]It is well-known from the literature on fixed-effects in panel models that $\hat{\psi}_{ML}$ may be affected by the incidental bias B^{inc}/n, which in our case arises from the uncertainty attached to estimating an increasing number of states $X_{1:T}$ as T grows. However, the states are estimated very accurately in multi-factor DTSMs and the incidental bias is therefore unlikely to be important for estimating DTSMs with a reasonable number of cross-section observations n. An analytical expression for the incidental bias B^{inc} may be derived following the procedure in Hahn and Newey (2004).

in equation (12), observe that

$$\begin{aligned}\frac{\partial L\left(\psi|Y_{1:T}\right)}{\partial X_t} &= -d_t R_w^{-1}\left(X_t - h\left(X_{t-1};\psi\right)\right) \\ &\quad + h_X\left(X_t;\psi\right)' R_w^{-1}\left(X_{t+1} - h\left(X_t;\psi\right)\right) \\ &\quad - g_X\left(X_t;\psi\right)' R_{\varepsilon,t}^{-1}\left(Y_t - g\left(X_t;\psi\right)\right) \\ &= 0\end{aligned} \tag{14}$$

for $t = 1, 2, ..., T$, where $d_1 = 0$ and $d_t = 1$ for $t > 1$. The matrix $h_X\left(X_t;\psi\right) \equiv \frac{\partial h(X_t;\psi)}{\partial X_t'}$ and denotes the $n_x \times n_x$ Jacobian of $h\left(X_t;\psi\right)$ with respect to X_t, and similarly, $g_X\left(X_t;\psi\right) \equiv \frac{\partial g(X_t;\psi)}{\partial X_t'}$ with dimensions $n_{y,t} \times n_x$. Let $X_{1:T}^{(i)}$ denote the points around which the system in equations (10) and (11) in the paper is linearized. That is,

$$Y_t = g\left(X_t^{(i)};\psi\right) + g_X\left(X_t^{(i)};\psi\right)\left(X_t - X_t^{(i)}\right) + \varepsilon_t, \tag{15}$$

$$X_{t+1} = h\left(X_t^{(i)};\psi\right) + h_X\left(X_t^{(i)};\psi\right)\left(X_t - X_t^{(i)}\right) + w_{t+1}. \tag{16}$$

For a given value of ψ, let $\hat{X}_{1:T}\left(\psi\right)$ denote the state estimates from running the Kalman smoother on the linearized system in equations (15) and (16). As shown in Durbin and Koopman (2001), $\hat{X}_{1:T}\left(\psi\right)$ then solves the following system of equations

$$\begin{aligned}&-d_t R_w^{-1}\left(X_t - h\left(X_{t-1}^{(i)};\psi\right) - h_X\left(X_{t-1}^{(i)};\psi\right)\left(\hat{X}_{t-1}\left(\psi\right) - X_{t-1}^{(i)}\right)\right) \\ &+ h_X\left(X_t^{(i)};\psi\right)' R_w^{-1}\left(X_{t+1} - h\left(X_t^{(i)};\psi\right) - h_X\left(X_t^{(i)};\psi\right)\left(\hat{X}_t\left(\psi\right) - X_t^{(i)}\right)\right) \\ &- g_X\left(X_t^{(i)};\psi\right)' R_{\varepsilon,t}^{-1}\left(Y_t - g\left(X_t^{(i)};\psi\right) - g_X\left(X_t^{(i)};\psi\right)\left(\hat{X}_t\left(\psi\right) - X_t^{(i)}\right)\right) \\ = \; & 0\end{aligned} \tag{17}$$

for $t = 1, 2, ..., T$. Accordingly, when $\hat{X}_t\left(\psi\right) = X_t^{(i)}\left(\psi\right)$ for all t, the conditions in equations (17) reduces to those in equations (14), meaning that $\hat{X}_{1:T}\left(\psi\right)$ is the ML estimates of the states for a given value of ψ. The iterative procedure to find this solution is as follows:

Step 1 Run the extended Kalman filter and smoother on the system in equations (10) and (11) to obtain $\hat{X}_{1:T}^{EKF}\left(\psi\right)$. Set $i = 1$ and let $X_{1:T}^{(i)}\left(\psi\right) = \hat{X}_{1:T}^{EKF}\left(\psi\right)$.

Step 2 Run the extended Kalman filter and smoother on the linearized system in equations (15) and (16) to obtain $\hat{X}_{1:T}\left(\psi\right)$.

Par.	QML		ML	
	Est	SE	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.1060	0.0763	0.0775	0.0774
$\kappa_{22}^{\mathbb{P}}$	0.2157	0.1443	0.2596	0.1585
$\kappa_{33}^{\mathbb{P}}$	0.7255	0.3649	0.8083	0.3275
σ_{11}	0.0052	0.0001	0.0052	0.0001
σ_{22}	0.0103	0.0010	0.0106	0.0006
σ_{33}	0.0207	0.0015	0.0219	0.0012
$\theta_1^{\mathbb{P}}$	0.0529	0.0034	0.0502	0.0102
$\theta_2^{\mathbb{P}}$	-0.0275	0.0093	-0.0258	0.0084
$\theta_3^{\mathbb{P}}$	-0.0230	0.0060	-0.0210	0.0060
λ	0.3747	0.0105	0.3738	0.0038

Table 13: **ML Estimates of the AFNS Model**
This table reports the estimated parameters (Est) in the AFNS model with independent factors in the one-step approach, using either QML or ML. The standard errors (SE) for the QML are computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, computed as outlined in Harvey (1989). The SE for the ML estimates are obtained as the inverse of the Hessian matrix for the concentrated log-likelihood function. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

Step 3 If $\left|\hat{X}_t(\psi) - X_t^{(i)}(\psi)\right| > \epsilon$ for any t, where ϵ is a small number, let $i = i+1$ and $X_{1:T}^{(i)}(\psi) = \hat{X}_{1:T}(\psi)$ and go to step 2, otherwise stop.

Let $\hat{X}_{1:T}(\psi)$ denote the states from this procedure, which depends on ψ. The concentrated log-likelihood function is then $L^c(\psi|Y_{1:T}) \equiv L\left(\psi, \hat{X}_{1:T}(\psi)|Y_{1:T}\right)$, which we optimize across ψ to obtain the ML estimates. The asymptotic standard errors for $\hat{\psi}_{ML}$ are obtained in a standard fashion, i.e.,

$$\widehat{AVar}\left(\hat{\psi}_{ML}\right) = \left[\sum_{t=1}^{T} s_i^c\left(\hat{\psi}_{ML}\right) s_i^c\left(\hat{\psi}_{ML}\right)'\right]^{-1}, \tag{18}$$

or the inverse of the concentrated Hessian matrix. Here, $s_i^c\left(\hat{\psi}_{ML}\right)$ denotes the concentrated score function computed by numerical differentiation of $L^c(\psi|Y_{1:T})$.

F.1 Results

The ML estimates are provided in Table 13 and seen to be very similar to the QML estimates based on the EKF. In particular, the two estimates of λ are almost identical. We also note that the standard errors for most parameters are smaller for the ML estimates when compared

Maturity bucket in years	No. obs.	QML		ML	
		Mean	RMSE	Mean	RMSE
0-2	1,472	-0.09	5.75	-0.11	5.81
2-4	1,098	0.46	4.76	0.40	4.79
4-6	744	-0.32	4.01	-0.47	3.90
6-8	404	-1.24	5.46	-1.38	5.49
8-10	477	-2.54	6.07	-2.70	6.13
10-12	289	-1.07	6.36	-1.31	6.24
12-14	155	3.78	6.72	3.56	6.54
14-16	168	0.76	4.32	0.61	4.27
16-18	179	0.70	4.66	0.59	4.54
18-20	192	1.71	4.33	1.51	4.19
20-22	186	3.64	4.98	3.07	4.47
22-24	142	0.84	4.74	0.21	4.04
24-26	124	0.05	5.63	-0.46	4.92
26-28	113	-5.45	8.71	-5.96	8.46
28<	288	-4.97	11.98	-5.32	12.46
All bonds	6,031	-0.33	5.90	-0.48	5.81

Table 14: **Summary Statistics of Bond Fitted Errors in the AFNS Model Estimated with ML**

This table reports the mean pricing errors (Mean) and the root mean-squared pricing errors (RMSE) of the Canadian bond prices for the AFNS model with independent factors estimated with the one-step approach using QML and ML as described in the main text. The pricing errors are reported in basis points and computed as the difference between the implied yield on the coupon bond and the model-implied yield on this bond. The data are monthly and cover the period from January 31, 2000, to April 29, 2016.

to the QML estimates. The close similarlity between the two set of estimates is also seen from the goodness of fit, as the ML estimates give an overall RMSE of 5.81 basis points, and hence basically provides the same satisfying fit as the EKF. In the paper, we therefore only report results using the QML estimator, which is computationally much faster to implement than the ML estimation described above.

G Full ML Estimation with Stochastic Volatility

Suppose X_t evolves as

$$X_{t+1} = h\left(X_t;\theta\right) + f\left(X_t, \varepsilon_{t+1};\theta\right), \tag{19}$$

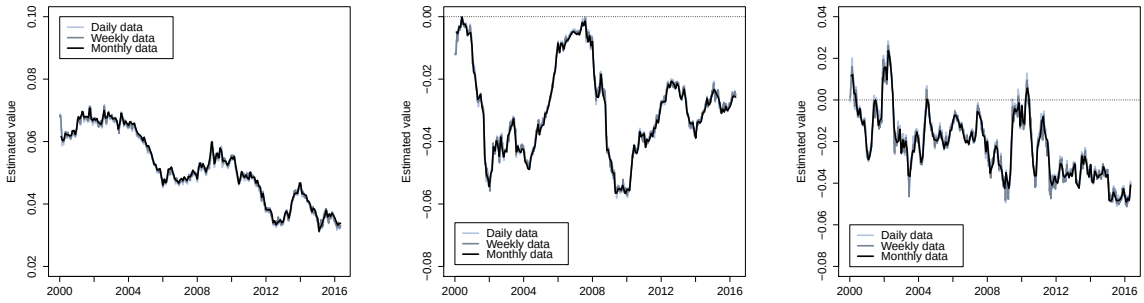
where the f-function accommodates stochastic volatility in the dynamics of X_t. Lagging equation (19) by one period and inserting it into equation (10) gives

$$\begin{aligned} Y_t &= \underbrace{g\left(h\left(X_{t-1};\theta\right) + f\left(X_{t-1},\varepsilon_t\right);\theta\right)}_{\tilde{g}(X_{t-1},\varepsilon_t;\theta)} + v_t \qquad (20) \\ \begin{bmatrix} X_t \\ \varepsilon_{t+1} \end{bmatrix} &= \begin{bmatrix} h\left(X_{t-1};\theta\right) + f\left(X_{t-1},\varepsilon_t\right) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ u_{t+1} \end{bmatrix} \end{aligned}$$

where $u_{t+1} \sim \mathcal{NID}\left(0, R_\varepsilon\right)$. Thus, by expanding the state vector to $\tilde{X}_t = \begin{bmatrix} X_{t-1}' & \varepsilon_t' \end{bmatrix}'$, the law of motion in equation (19) with nonlinearities between the states and the innovations can be rewritten into an extended system with only linear innovations as in equations (10) and (11).

H Sensitivity of One-Step Approach to Data Frequency

In this appendix, we document that the reported results for the benchmark AFNS model estimated with our preferred one-step approach using pricing errors scaled by duration are robust to varying the time frequency of the bond price data. Specifically, we repeat the estimations using daily and weekly (Fridays) data instead of the end-of-month data considered in the paper.



(a) L_t (b) S_t (c) C_t

Figure 4: **Estimated States in the AFNS Model for Different Data Frequencies**
Illustration of the estimated level, slope, and curvature factors in the AFNS model with independent factor dynamics estimated with the one-step approach using QML and different data frequencies.

Par.	Daily data		Weekly data		Monthly data	
	Est	SE	Est	SE	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.0173	0.0167	0.0117	0.0112	0.1060	0.0763
$\kappa_{22}^{\mathbb{P}}$	0.1301	0.0674	0.1782	0.1015	0.2157	0.1443
$\kappa_{33}^{\mathbb{P}}$	0.6182	0.1986	0.7006	0.2869	0.7255	0.3649
σ_{11}	0.0052	0.0000	0.0051	0.0001	0.0052	0.0001
σ_{22}	0.0071	0.0002	0.0084	0.0005	0.0103	0.0010
σ_{33}	0.0182	0.0005	0.0198	0.0009	0.0207	0.0015
$\theta_1^{\mathbb{P}}$	0.0730	0.0018	0.0797	0.0028	0.0529	0.0034
$\theta_2^{\mathbb{P}}$	-0.0227	0.0070	-0.0246	0.0086	-0.0275	0.0093
$\theta_3^{\mathbb{P}}$	-0.0209	0.0053	-0.0212	0.0063	-0.0230	0.0060
λ	0.3804	0.0023	0.3832	0.0050	0.3747	0.0105

Table 15: **Estimates of the AFNS Model with Different Data Frequencies**
This table reports the estimated parameters (Est) in the AFNS model with independent factors in the one-step approach using QML and different data frequencies. The standard errors (SE) are computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, as outlined in Harvey (1989).

Figure 4 compares the estimated paths of each of the three state variables, which are all

practically indistinguishable.

A similar closeness in results is observed in Table 15, which compares the estimated dynamic parameters. Although there is a tendency to greater persistence and lower volatility of the state variables as we increase the data frequency, we find the results overall to be very similar, in particular when the parameter uncertainty is taken into consideration.

These observations explain why we choose to use monthly data in the paper as it reduces the computing time significantly. However, we stress that all our exercises could have been performed with weekly or even daily data.

I Formulas for Short Rate Expectations and Term Premiums

In this appendix, we derive the analytical formulas for short-rate expectations and term premiums in the independent-factor AFNS model. Recall that the term premium is defined as

$$TP_t(\tau) = y_t(\tau) - \frac{1}{\tau}\int_t^{t+\tau} E_t^{\mathbb{P}}[r_s]ds.$$

Furthermore, recall that for affine stochastic differential equations of the form

$$dX_t = \mathcal{K}^{\mathbb{P}}(\theta^{\mathbb{P}} - X_t)dt + \Sigma dW_t^{\mathbb{P}},$$

the conditional expectation is given by

$$E_t^{\mathbb{P}}[X_{t+\tau}] = (I - \exp(-\mathcal{K}^{\mathbb{P}}\tau))\theta^{\mathbb{P}} + \exp(-\mathcal{K}^{\mathbb{P}}\tau)X_t.$$

In the AFNS model, the instantaneous short rate is defined as

$$r_t = L_t + S_t,$$

while the specification of the $\mathbb{P}$-dynamics considered is given by

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{22}^{\mathbb{P}} & 0 \\ 0 & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left[\begin{pmatrix} \theta_1^{\mathbb{P}} \\ \theta_2^{\mathbb{P}} \\ \theta_3^{\mathbb{P}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{P}} \\ dW_t^{S,\mathbb{P}} \\ dW_t^{C,\mathbb{P}} \end{pmatrix}.$$

Thus, the mean-reversion matrix is given by

$$\mathcal{K}^{\mathbb{P}} = \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{22}^{\mathbb{P}} & 0 \\ 0 & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix}.$$

Its matrix exponential can be calculated analytically and is given by

$$\exp(-\mathcal{K}^{\mathbb{P}}\tau) = \begin{pmatrix} e^{-\kappa_{11}^{\mathbb{P}}\tau} & 0 & 0 \\ 0 & e^{-\kappa_{22}^{\mathbb{P}}\tau} & 0 \\ 0 & 0 & e^{-\kappa_{33}^{\mathbb{P}}\tau} \end{pmatrix}.$$

Now, the conditional mean of the state variables is

$$E_t^{\mathbb{P}}[X_{t+\tau}] = \theta^{\mathbb{P}} + \begin{pmatrix} e^{-\kappa_{11}^{\mathbb{P}}\tau} & 0 & 0 \\ 0 & e^{-\kappa_{22}^{\mathbb{P}}\tau} & 0 \\ 0 & 0 & e^{-\kappa_{33}^{\mathbb{P}}\tau} \end{pmatrix} \begin{pmatrix} L_t - \theta_1^{\mathbb{P}} \\ S_t - \theta_2^{\mathbb{P}} \\ C_t - \theta_3^{\mathbb{P}} \end{pmatrix} = \begin{pmatrix} \theta_1^{\mathbb{P}} + e^{-\kappa_{11}^{\mathbb{P}}\tau}(L_t - \theta_1^{\mathbb{P}}) \\ \theta_2^{\mathbb{P}} + e^{-\kappa_{22}^{\mathbb{P}}\tau}(S_t - \theta_2^{\mathbb{P}}) \\ \theta_3^{\mathbb{P}} + e^{-\kappa_{33}^{\mathbb{P}}\tau}(C_t - \theta_3^{\mathbb{P}}) \end{pmatrix}.$$

In order to get back to the term premium formula, we note that the conditional expectation of the instantaneous short rate process is:

$$\begin{aligned} E_t^{\mathbb{P}}[r_s] &= E_t^{\mathbb{P}}[L_s + S_s] \\ &= \theta_1^{\mathbb{P}} + e^{-\kappa_{11}^{\mathbb{P}}(s-t)}(L_t - \theta_1^{\mathbb{P}}) + \theta_2^{\mathbb{P}} + e^{-\kappa_{22}^{\mathbb{P}}(s-t)}(S_t - \theta_2^{\mathbb{P}}). \end{aligned}$$

Next, we integrate the expected short rate over the time interval from t to $t+\tau$ as in the definition of the term premium:

$$\begin{aligned} \int_t^{t+\tau} E_t^{\mathbb{P}}[r_s]ds &= \int_t^{t+\tau} \Big(\theta_1^{\mathbb{P}} + e^{-\kappa_{11}^{\mathbb{P}}(s-t)}(L_t - \theta_1^{\mathbb{P}}) + \theta_2^{\mathbb{P}} + e^{-\kappa_{22}^{\mathbb{P}}(s-t)}(S_t - \theta_2^{\mathbb{P}})\Big)ds \\ &= (\theta_1^{\mathbb{P}} + \theta_2^{\mathbb{P}})\tau + (L_t - \theta_1^{\mathbb{P}})\frac{1 - e^{-\kappa_{11}^{\mathbb{P}}\tau}}{\kappa_{11}^{\mathbb{P}}} + (S_t - \theta_2^{\mathbb{P}})\frac{1 - e^{-\kappa_{22}^{\mathbb{P}}\tau}}{\kappa_{22}^{\mathbb{P}}}. \end{aligned}$$

The relevant term to go into the term premium formula is the average expected short rate

$$\frac{1}{\tau}\int_t^{t+\tau} E_t^{\mathbb{P}}[r_s]ds = \theta_1^{\mathbb{P}} + \theta_2^{\mathbb{P}} + (L_t - \theta_1^{\mathbb{P}})\frac{1 - e^{-\kappa_{11}^{\mathbb{P}}\tau}}{\kappa_{11}^{\mathbb{P}}\tau} + (S_t - \theta_2^{\mathbb{P}})\frac{1 - e^{-\kappa_{22}^{\mathbb{P}}\tau}}{\kappa_{22}^{\mathbb{P}}\tau}.$$

The final expression for the term premium is then given by

$$\begin{aligned} TP_t(\tau) &= y_t(\tau) - \frac{1}{\tau}\int_t^{t+\tau} E_t^{\mathbb{P}}[r_s]ds \\ &= L_t + \frac{1 - e^{-\lambda\tau}}{\lambda\tau}S_t + \Big(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\Big)C_t - \frac{A(\tau)}{\tau} \\ &\quad -\theta_1^{\mathbb{P}} - \theta_2^{\mathbb{P}} - (L_t - \theta_1^{\mathbb{P}})\frac{1 - e^{-\kappa_{11}^{\mathbb{P}}\tau}}{\kappa_{11}^{\mathbb{P}}\tau} - (S_t - \theta_2^{\mathbb{P}})\frac{1 - e^{-\kappa_{22}^{\mathbb{P}}\tau}}{\kappa_{22}^{\mathbb{P}}\tau} \\ &= \Big(1 - \frac{1 - e^{-\kappa_{11}^{\mathbb{P}}\tau}}{\kappa_{11}^{\mathbb{P}}\tau}\Big)L_t + \Big(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - \frac{1 - e^{-\kappa_{22}^{\mathbb{P}}\tau}}{\kappa_{22}^{\mathbb{P}}\tau}\Big)S_t + \Big(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\Big)C_t \\ &\quad -\Big(1 - \frac{1 - e^{-\kappa_{11}^{\mathbb{P}}\tau}}{\kappa_{11}^{\mathbb{P}}\tau}\Big)\theta_1^{\mathbb{P}} - \Big(1 - \frac{1 - e^{-\kappa_{22}^{\mathbb{P}}\tau}}{\kappa_{22}^{\mathbb{P}}\tau}\Big)\theta_2^{\mathbb{P}} - \frac{A(\tau)}{\tau}. \end{aligned}$$

J Details of the Forecast Exercise

In this appendix, we provide the full details of the forecast exercise in Section 5 of the paper.

J.1 The Consensus Forecasts

We use the Consensus Forecasts survey as the benchmark in our forecast exercise for at least three reasons. First and most importantly, it offers a long history of forecasts of Canadian bond yields. Second, the fixed structure of its survey questions is particularly suitable for a real-time forecast exercise like ours as we explain below. Finally, we note that it tracks a panel of very qualified economic forecasters. To give an example, the May 2016 survey, which is the last survey used in our exercise, included interest rate projections from a total of 15 participating institutions. Thus, we consider these forecasts to be reliable and of high quality. As a consequence, they serve as a good yardstick for validating the performance of the various models and estimation approaches we consider.

The survey is performed once a month, and participants are asked to submit forecasts for two Canadian interest rates, the three-month Treasury bill rate and ten-year government bond yield, at two forecast horizons, namely at the end of the third month after the survey month and at the end of the survey month the following year. Since the survey dates are typically the second Monday of the survey month, this structure implies that the effective forecast horizons are roughly three and a half months and twelve and a half months. However, for convenience we refer to them as three- and twelve-month forecasts, respectively, although we stress that we generate the model-implied forecast to match exactly the future dates indicated in each survey. It is this repeated regular pattern to the survey questions that makes it well-suited as a benchmark in a real-time forecast exercise.

Since we only consider Canadian government bond price data back to January 2000, there is a limitation on how early we can start the forecast exercise. Due to the high persistence of the state variables, the models need some minimum sample length to be able to appropriately estimate the factor dynamics. At the same time, we want the out-of-sample forecast period to be as long as possible to increase the confidence of the inferences that we try to draw about the model forecast performance. To balance this tradeoff, we choose to start the forecast exercise at the end of December 2006 for the real-time model estimations. This choice implies that the first survey forecasts we consider are from the Consensus Forecasts survey dated

January 8, 2007, and we focus on the consensus forecasts, i.e., the mean of the individual forecasts for the four (yield, forecast horizon)-pairs available in each survey.

J.2 Yield Forecast Generation

To explain the matching yield forecast generation from the estimated DTSMs, consider the ith survey dated t_0^i with the two forecast dates T_1^i and T_2^i. To map this to our models, we use the estimated model parameters from the bond price data up until the end of the month before the survey month, denoted t^i with $t^i < t_0^i$. This means that the model forecasts are lagged by about 10 days relative to the survey date of the economic forecasters. This makes our forecast performance assessment conservative relative to the survey panel. This also implies that the two effective forecast horizons for the model forecasts are $\Delta_1^i = T_1^i - t^i$ and $\Delta_2^i = T_2^i - t^i$.

For the forecasts of the three-month Treasury bill rate, we treat them as forecasts of three-month zero-coupon yields.

In the AFNS and AFGNS models, this yield is affine in the state variables

$$y_t(3m) = A(3m) + B(3m)' X_t.$$

The conditional expectation of the state variables is easily calculated as

$$E_{t^i}^{\mathbb{P}}[X_{t^i+\Delta_j^i}] = (I - \exp(-\widehat{\mathcal{K}}_{t^i}^{\mathbb{P}} \Delta_j^i))\widehat{\theta}_{t^i}^{\mathbb{P}} + \exp(-\mathcal{K}^{\mathbb{P}} \Delta_j^i) X_{t^i} \quad \text{for} \quad j = 1, 2,$$

where $\widehat{\mathcal{K}}_{t^i}^{\mathbb{P}}$ and $\widehat{\theta}_{t^i}^{\mathbb{P}}$ are the estimated model parameters using bond price data up until date t^i.

Hence, the forecasts from these two models for the three-month yield corresponding to the ith survey are given by

$$E_{t^i}^{\mathbb{P}}[y_{t^i+\Delta_j^i}(3m)] = A(3m) + B(3m)' E_{t^i}^{\mathbb{P}}[X_{t^i+\Delta_j^i}] \quad \text{for} \quad j = 1, 2.$$

In the B-AFNS model, even zero-coupon yields are not linear functions of the state variables. As a consequence, we have to resort to Monte Carlo simulations to generate the forecasts for the three-month yields from this model.

To do so, we first calculate the estimated conditional covariance matrix of the state variables

$$\widehat{Q}_j^i = \int_0^{\Delta_j^i} e^{-\widehat{\mathcal{K}}_{t^i}^{\mathbb{P}} s} \widehat{\Sigma}_{t^i} \widehat{\Sigma}_{t^i}' e^{-(\widehat{\mathcal{K}}_{t^i}^{\mathbb{P}})' s} ds \quad \text{for} \quad j = 1, 2$$

using the estimated $\widehat{\mathcal{K}}^{\mathbb{P}}_{t^i}$ and $\widehat{\Sigma}_{t^i}$ matrices as of date t^i.

Now, let $\widehat{\Sigma}^i_j$ be the Cholesky decomposition of $\widehat{Q}^i_j$ and let $Z(\omega)$ be a vector of random variables that are each $\mathcal{N}(0,1)$ distributed with dimension equal to the number of state variables in the model. Then

$$X_{t^i+\Delta^i_j}(\omega) = (I - \exp(-\widehat{\mathcal{K}}^{\mathbb{P}}_{t^i}\Delta^i_j))\widehat{\theta}^{\mathbb{P}}_{t^i} + \exp(-\mathcal{K}^{\mathbb{P}}\Delta^i_j)X_{t^i} + \widehat{\Sigma}^i_j Z(\omega)$$

represents a random draw of the state variables at time T^i_j for $j = 1, 2$.

Next, we repeat this $N = 5{,}000$ times to get draws $X_{t^i+\Delta^i_j}(\omega^n)$ for $n = 1, \ldots, N$. The yield forecast is then given by the mean of the value of the nonlinear yield function $g(X_t)$ evaluated at each of the draws:

$$E^{\mathbb{P}}_{t^i}[y_{t^i+\Delta^i_j}(3m)] = \frac{1}{N}\sum_{n=1}^{N} g(X_{t^i+\Delta^i_j}(\omega^n)) \quad \text{for} \quad j = 1, 2.$$

For the forecasts of the ten-year government bond yield, we note that the Bank of Canada (just like the U.S. Treasury) tends to issue new bonds as close to par as possible (subject to minimum increments of 12.5 basis points in the stated coupon rate). We take this to mean that the survey participants are projecting ten-year par-coupon yields even though the Bank of Canada only issues new ten-year bonds roughly once a year. Thus, at times it may be known in advance (in particular for the three-month forecasts) that there may not be any new ten-year bonds trading, but only old seasoned bonds that could be trading some distance away from par. Even under those circumstances we take the submitted forecasts to represent forecasts of ten-year par-coupon yields. By implication, we must generate forecasts of ten-year par-coupon yields from the models.

To begin, we note that synthetic ten-year par-coupon yields are calculated by adjusting the coupon rate C in the following equation

$$1 = \sum_{j=1}^{20} \frac{C}{2} \exp\left\{-\widehat{y}_{t^i}\left(t_j - t\right)\right\} + \exp\left\{-\widehat{y}_{t^i}\left(10\right)\right\}, \tag{21}$$

where $\widehat{y}_{t^i}\tau$ is the fitted τ-year zero-coupon yield implied by the considered model estimated as of time t^i.

Since this is a nonlinear function in the state variables, we again have to resort to Monte Carlo simulation of the state variables identical to the algorithm described above. For each

survey date t^i and forecast horizon T^i_j, this gives us N estimated ten-year par-coupon rates denoted $C^i_j(\omega^n)$, $n = 1, \ldots, N$.

The forecast of the ten-year par-coupon yield is then the average of the individual projected par-coupon values:

$$E^{\mathbb{P}}_{t^i}[C^i_j] = \frac{1}{N}\sum_{n=1}^{N} C^i_j(\omega^n) \quad \text{for} \quad j = 1, 2.$$

Finally, this is repeated for all models and all $i = 1, \ldots, I$ survey dates, where $I = 113$ since we cover the period from end of December 2006 to end of April 2016 with the matching Consensus forecasts covering the period from January 8, 2007, to May 9, 2016.

J.3 Yield Realizations

When it comes to the yield realizations, the available data forces us to use two different approaches.

For the three-month Treasury bill rates, we linearly interpolate between the end of the month readings of the rates of the two Treasury bills whose remaining times to maturity provide the tightest bracket around the three-month maturity point that we use in the generation of the model forecasts and that the participants in the Consensus Forecasts survey panel are assumed to be predicting.

As for the ten-year government bond yields, we base the generated forecast on par-coupon yields as explained in the previous section. Unfortunately, as already noted, the Bank of Canada only issues new ten-year bonds that would be trading close to par rather infrequently (roughly once a year). As a consequence, we have to estimate what the coupon rate would be on a hypothetical new ten-year bond issued at par at the end of each month included in our forecast exercise.

To generate these realizations, we exploit the finding in Section 5.2 of the paper that the AFGNS model delivers an accurate fit to the universe of Canadian government bonds. We therefore estimate the AFGNS model on an updated sample of Canadian government bond prices that contain data through the end of December 2017.

Figure 5 shows the yields to maturity through December 2017 for all 118 bonds in the updated sample, while Table 16 reports the estimated parameters for the AFGNS model using the updated sample of bond prices.

We next combine the estimated parameters and state variables to calculate synthetic ten-

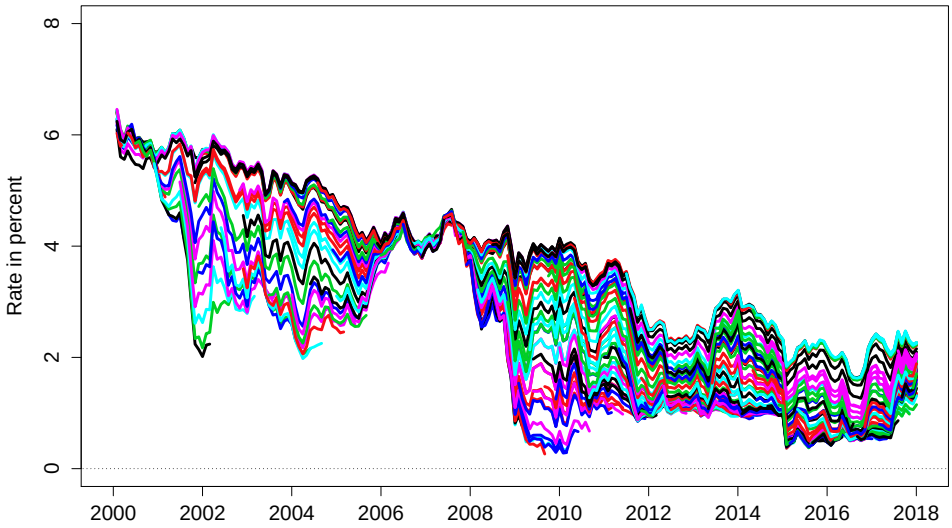


Figure 5: **Updated Sample of Canadian Bond Yields**

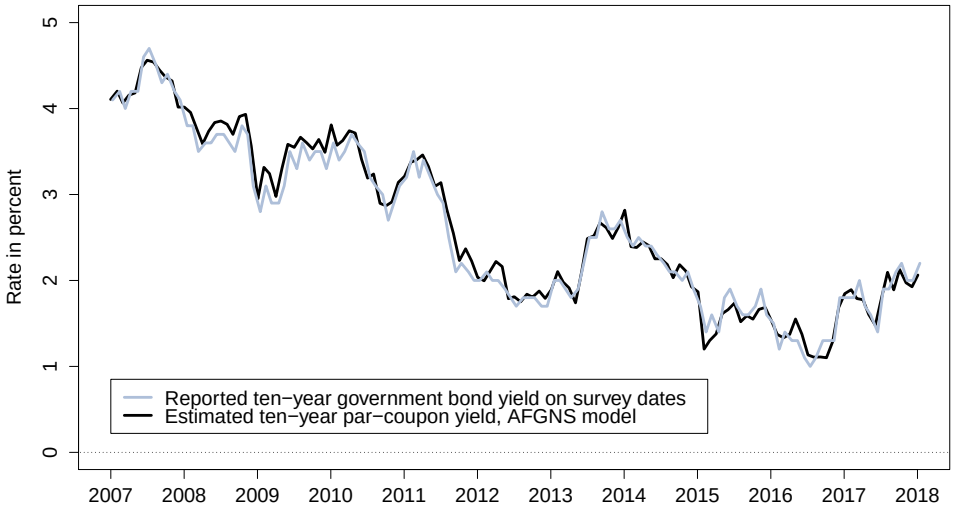


Figure 6: **Estimated Ten-Year Par-Coupon Bond Yield**

Par.	One-step approach	
	Est	SE
$\kappa_{11}^{\mathbb{P}}$	0.0445	0.0467
$\kappa_{22}^{\mathbb{P}}$	0.2129	0.1508
$\kappa_{33}^{\mathbb{P}}$	0.2618	0.2123
$\kappa_{44}^{\mathbb{P}}$	0.7314	0.3037
$\kappa_{55}^{\mathbb{P}}$	0.2024	0.1398
σ_{11}	0.0028	0.0005
σ_{22}	0.0123	0.0009
σ_{33}	0.0106	0.0009
σ_{44}	0.0225	0.0019
σ_{55}	0.0208	0.0012
$\theta_1^{\mathbb{P}}$	0.0517	0.0049
$\theta_2^{\mathbb{P}}$	0.0137	0.0133
$\theta_3^{\mathbb{P}}$	-0.0417	0.0104
$\theta_4^{\mathbb{P}}$	0.0079	0.0079
$\theta_5^{\mathbb{P}}$	0.0396	0.0157
λ	0.6594	0.0272
$\tilde{\lambda}$	0.1202	0.0070

Table 16: **Parameter Estimates in the Updated AFGNS Model**
This table reports the estimated parameters (Est) in the AFGNS model with independent factors and their standard errors (SE) using either the one-step. The SE approach are computed by pre- and post-multiplying the variance of the score by the inverse of the Hessian matrix, as outlined in Harvey (1989). The data are monthly and cover the period from January 31, 2000, to December 29, 2017.

year par-coupon yields based on equation (21) starting a the end of December 2006. This produces the time series shown in Figure 6, which are the values we use as the realizations in calculating the errors for the ten-year par-coupon yield forecasts.

To validate the accuracy of the series, we compare them to the ten-year yields on the survey dates as reported in the Consensus Forecasts surveys and also shown in Figure 6. We note the closeness of the two yield series, which offers support for our approach. We also note two reasons why they are not perfectly aligned. First and most importantly, they are observed on two different dates each month (last day of the month versus the Consensus Forecasts survey date). Second, the numbers reported in the surveys are rounded to the first decimal, whereas our model-implied series are reported with 7 decimals accuracy. Both of these differences are likely to contribute to the time-varying wedges between the two series.

References

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015, "Parameter Inference and Dynamic State Recovery from Option Panels," *Econometrica*, Vol 83, No. 3, 1081-1145.

Andreasen, Martin M., 2013, "Non-Linear DSGE Models and The Central Difference Kalman Filter," *Journal of Applied Econometrics*, Vol. 28, 929-955.

Andreasen, Martin M. and Bent Jesper Christensen, 2015, "The SR Approach: A New Estimation Procedure for Non-Linear and Non-Gaussian Dynamic Term Structure Models," *Journal of Econometrics*, Vol. 184, No. 2, 420-451.

Bai, Jushan, 2003, "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, Vol. 71, No. 1, 135–171.

Bai, Jushan, and Serena Ng, 2002, "Determining the Number of Factors in Approximate Factor Models,"*Econometrica*, Vol. 70, No. 1, 191–221.

Bollerslev, Tim and Jeffrey M. Wooldridge, 1992, "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances," *Econometrics Reviews*, Vol. 11, No. 2, 143-172.

Christensen, Jens H. E. and Glenn D. Rudebusch, 2015, "Estimating Shadow-Rate Term Structure Models with Near-Zero Yields," *Journal of Financial Econometrics*, Vol. 13, No. 2, 226-259.

Durbin, J. and S. J. Koopman, 2001, "Time Series Analysis by State Space Methods," Oxford University Press, Oxford.

Filipović, Damir and Anders B. Trolle, 2013, "The Term Structure of Interbank Risk," *Journal of Financial Economics*, Vol. 109, No. 3, 707-733.

Fisher, Mark and Christian Gilles, 1996, "Term Premia in Exponential-Affine Models of the Term Structure," Manuscript, Board of Governors of the Federal Reserve System.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2007, "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, Vol. 54, No. 8, 2291-2304.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2010, "The TIPS Yield Curve and Inflation Compensation," *American Economic Journal: Macroeconomics*, Vol. 2, No. 1, 70-92.

Hahn, Jinyong and Whitney Newey, 2004, "Jackknife and Analytical Bias Reduction for Nonlinear Panel Models," *Econometrica*, Vol. 72, No. 4, 1295-1319.

Harvey, A.C., 1989, *Forecasting, structural time series models and the Kalman filter,* Cambridge: Cambridge University Press.

Nelson, Charles R. and Andrew F. Siegel, 1987, "Parsimonious Modeling of Yield Curves," *Journal of Business*, Vol. 60, No. 4, 473-489.

Pancost, N. Aaron, 2018, "Zero-Coupon Yields and the Cross-Section of Bond Prices," *Chicago University, Working Paper*, 1-66.

Svensson, Lars E. O., 1995, "Estimating Forward Interest Rates with the Extended Nelson-Siegel Method," *Quarterly Review*, No. 3, Sveriges Riksbank, 13-26.