中央大学企業研究所
Working Paper No.51

How Arbitrage-Free is the Nelson-Siegel Model under Stochastic Volatility?

Hideyuki Takamizawa

May, 2019

Working Paper Series

THE INSTITUTE OF BUSINESS RESEARCH
Chuo University
Tokyo, Japan

How Arbitrage-Free is the Nelson-Siegel Model under Stochastic Volatility?

Hideyuki Takamizawa *

Faculty of Commerce, Chuo University
Hachioji Tokyo 192-0393, JAPAN
E-mail: takamizawa@tamacc.chuo-u.ac.jp

First draft: February 7, 2019
This draft: May 6, 2019

Abstract

This study examines the effect of no-arbitrage on the Nelson-Siegel (NS) yield curve under stochastic interest-rate volatility. Unlike under constant volatility, in which only a constant (convexity-adjustment) term of a yield function differs with and without no-arbitrage, factor loadings also differ when the volatility is spanned by interest-rate factors. After controlling for the drift, we find that spanned volatility does not magnify the effect of no-arbitrage relative to constant volatility. The finding supports a conventional use of the NS model such that the model is augmented with stochastic volatility after identifying the factors from the yield curve.

Keywords: Nelson-Siegel model, Yield curve, Interest rate, No-arbitrage, Stochastic volatility.

*This research is supported by JSPS KAKENHI (Grant Number 15K03538).

1 Introduction

This study examines the effect of no-arbitrage on the Nelson and Siegel (NS) (1987) model when the volatility of changes in interest-rate factors is stochastic. The NS model is a convenient tool for describing the cross-section of interest rates, and hence used for both practical and academic purposes.[1] Diebold and Li (DL) (2006) augment the NS model with dynamics. Specifically, they interpret the loadings of the NS model that depend on the time to maturity as the explanatory variables, and the coefficients of the loadings as the factors. Then, they obtain the factors at each point in time by fitting the NS yield curve to the observed curve and estimate the dynamics of the factors from the pooled data. The dynamic NS model has also become a popular tool for studying both the time-series and cross-section of interest rates, such as in works by Diebold, Rudebusch, and Aruoba (2006), Diebold, Li, and Yue (2008), and Koopman, Mallee, and Van Der Wel (2010).

In contrast, there is a theoretical criticism that the NS model does not preclude arbitrage opportunities; see Björk and Christensen (1999), and Filipović (1999). Christensen, Diebold, and Rudebusch (CDR) (2009, 2011) present a solution that makes the dynamic NS model consistent with no-arbitrage, while preserving the same factor loadings as in the original NS model. The arbitrage-free NS model is a special case of affine term structure models. Due to the convexity adjustment, the no-arbitrage yield deviates from the original NS yield by a maturity-dependent constant. Hence, the magnitude of this constant indicates how large the effect of no-arbitrage is on the NS yield curve. Indeed, Coroneo, Nyholm, and Vidova-Koleva (2011) examine this difference and find it minor.[2] Chen and Du (2013) extend this argument to the case in which interest-rate volatility is driven by an unspanned factor; that is, a factor that does not affect the cross-section of interest rates. They employ the Heath, Jarrow, and Morton (1992) framework, with which it is straightforward to make a volatility factor unspanned; see Collin-Dufresne and Goldstein (2002). Since the loading of the yield function does not change when including the unspanned volatility factor, the argument reduces to the one with constant volatility, in which the difference in the model-implied yields with and without no-arbitrage is summarized in the convexity-adjustment term.

[1] Söderlind and Svensson (1997) extend the NS model by adding an additional hump term to improve the fit of the long end of the yield curve.

[2] Coroneo et al. (2011) do not follow CDR when specifying the risk neutral drift of changes in factors. Therefore, the factor loadings consistent with no-arbitrage are not exactly the same as those in the NS model.

Yet unanswered is the case in which interest-rate factors drive the volatility; that is, spanned volatility. Since the original NS model is silent about the volatility, or more generally the dynamics, there is no reason to exclude spanned volatility. In this case, not only the convexity-adjustment term but also the factor loadings may be different from those in the NS model.

The purpose of this study is to examine, under stochastic volatility spanned by interest-rate factors, the effect of no-arbitrage on the NS yield curve. In particular, we ask whether the extension from constant to spanned volatility magnifies the effect. Our examination relies on specific models and datasets because while the difference in yields with and without no-arbitrage does exist in theory, its magnitude is a practical issue, which requires some setup of the specification and estimation/calibration. However, given our research purpose, the choice of setup may not be an issue. Specifically, we aim to find an *incremental* effect from constant to spanned volatility, which does not seem to change from one setup to another because the effects with constant and spanned volatility are detected under the same conditions and then relativized.

To obtain a spanned-volatility model, we extend the arbitrage-free NS model proposed by CDR such that the covariance matrix of changes in factors depends on the level of factors while holding the risk neutral-drift unchanged. This extension enables us to detect an incremental effect of no-arbitrage on the NS yield curve attributed solely to spanned volatility. We consider two approaches to the detection.

The first approach focuses on the yield curve. Specifically, we feed both the factor and parameter values obtained without imposing no-arbitrage into the no-arbitrage version of the models. Then, we compute the ex-post no-arbitrage yield and compare it to the original NS yield. The second approach looks to the factors and parameters. In this approach, we extract the factors and estimate their dynamics simultaneously in both cases with and without no-arbitrage. Since the model-implied yields match the observed yields directly, they would differ little with and without no-arbitrage. Instead, the difference would arise in the extracted factors and estimated parameters.

We employ the dataset from DL to estimate the models, which allows us to compare our results for spanned volatility to those for constant volatility reported by Coroneo et al. (2010), who also use the DL dataset. To check the robustness, we also use an alternative dataset constructed by Gürkaynak, Sack, and Wright (GSW) (2007), which covers more recent observations.

We find from the first approach that the ex-post no-arbitrage yield deviates from the NS yield, but that the degree of deviation is similar on average between the constant- and spanned-volatility models. The deviation for a spanned-volatility model varies over time, depending on the level factor of interest rates. Specifically, when interest rates are high (low), the deviation becomes large (small). This pattern emerges because the model-implied yield has a constant term that increases with maturity and a loading on the level factor that decreases with maturity.

We find from the second approach that the volatility of the level factor is estimated lower when no-arbitrage is imposed than when it is not. However, the degree of decline in volatility is similar on average for the constant- and spanned-volatility models. Additionally, the spanned volatility does not magnify the difference in the extracted factors with and without no-arbitrage.

Taken together, the imposition of no-arbitrage alters the cross-section and/or time-series of interest rates implied originally by the (dynamic) NS model. However, the magnitude of these changes is not sensitive to the modeling of volatility, which has an important implication. If we can accept the inconsistency with no-arbitrage that the standard (dynamic) NS model induces, we can enrich the model with stochastic volatility, which might be specified independently of the extraction of factors.

The rest of the manuscript proceeds as follows. Section 2 presents the models. Sections 3 and 4 explain the data and estimation method, respectively. Section 5 presents the results and Section 6 concludes.

2 Model

We first present the dynamic NS model with constant volatility and its no-arbitrage version in Section 2.1, and then extend them with spanned volatility in Section 2.2.

2.1 The dynamic NS model and its no-arbitrage version

Following DL, let $X_t = (x_{1,t},\ x_{2,t},\ x_{3,t})'$ be a vector of factors driving the yield curve. Using X_t, a τ-maturity yield of the NS model is

$$Y^{NS}(X_t, \tau) = B(\tau)' X_t \ , \tag{1}$$

where $B(\tau)$ is a three-dimensional vector,

$$B(\tau) = \left(1, \qquad \frac{1 - e^{-\lambda\tau}}{\lambda\tau}, \qquad \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)' \ , \tag{2}$$

with $\lambda > 0$.

In the standard dynamic NS model considered by DL, the physical distribution of instantaneous changes in X_t is given by

$$dX_t \sim N[\ (K_0 + K_1 X_t)\ dt, \quad SRS\ dt\ ]\ , \tag{3}$$

where

$$K_0 = \begin{pmatrix} k_{0,1} \\ k_{0,2} \\ k_{0,3} \end{pmatrix}, \qquad K_1 = \begin{pmatrix} k_{1,11} & k_{1,12} & k_{1,13} \\ k_{1,21} & k_{1,22} & k_{1,23} \\ k_{1,31} & k_{1,32} & k_{1,33} \end{pmatrix}, \qquad R = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}, \tag{4}$$

$$S = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}, \tag{5}$$

with $|\rho_{ij}| < 1$ and $\sigma_i > 0$. We label the model consisting of (1) and (3) as "NS-CV" (Nelson-Siegel Constant-Volatility).

CDR make the dynamic NS model consistent with no-arbitrage. For this purpose, they specify the risk-neutral distribution of dX_t and the instantaneous risk-free rate r_t as

$$dX_t \sim N[\ K_1^Q X_t\ dt, \quad SRS\ dt\ ]\ , \tag{6}$$

$$r_t = x_{1,t} + x_{2,t}\ , \tag{7}$$

where

$$K_1^Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\lambda & \lambda \\ 0 & 0 & -\lambda \end{pmatrix}. \tag{8}$$

We label the model consisting of (3), (6), and (7) as "AF-CV" (Arbitrage-Free Constant-Volatility).

The no-arbitrage yield $Y^{AFCV}(X_t, \tau)$ is derived endogenously as

$$Y^{AFCV}(X_t, \tau) = A(\tau) + B(\tau)' X_t\ , \tag{9}$$

where $B(\tau)$ is given by (2) and

$$A(\tau) = -\frac{1}{2}\tau \int_0^\tau B(t)' SRSB(t) dt\ . \tag{10}$$

Comparing (9) with (1), we notice that the no-arbitrage yield differs from the NS yield only by a maturity-dependent constant, $A(\tau)$, which by (10) is the convexity-adjustment term, taking negative values for $\tau > 0$. Then, the magnitude of $A(\tau)$ measures the effect of no-arbitrage on the yield curve under constant volatility. To compute $A(\tau)$, the values of S and R (together with λ in $B(\tau)$) are needed, which this study estimates from the actual data.

When the volatility is spanned by interest-rate factors, it is not straightforward to examine the effect of no-arbitrage on the yield curve because both $A(\tau)$ and $B(\tau)$ may change from those in the NS model. In the next subsection, we consider a spanned-volatility model such that only the covariance matrix in the distribution of dX_t differs from that for the constant-volatility model.

2.2 An extended model with spanned volatility

We simply replace the constant covariance matrix in the NS/AF-CV model with a state-dependent one, and leave the other components of the model unchanged. This aims to control for the drift and address whether the spanned volatility magnifies the effect of no-arbitrage on the NS yield curve. Specifically, the physical and risk-neutral distributions of dX_t are given, respectively, by

$$dX_t \sim N[\ (K_0 + K_1 X_t)\, dt,\quad S_t R S_t\, dt\ ]\ , \tag{11}$$

$$dX_t \sim N[\ K_1^Q X_t\, dt,\quad S_t R S_t\, dt\ ]\ , \tag{12}$$

where K_0, K_1, K_1^Q, and R are the same as given by (4) and (8). As in the constant-volatility model, S_t is a diagonal matrix with the i-th diagonal element in absolute value interpreted as the standard deviation of the i-th factor. We specify the i-th diagonal element as a linear function of X_t: [3]

$$s_i(X_t) = \sigma_i + \beta_i' X_t \qquad (i = 1, 2, 3)\ . \tag{13}$$

Takamizawa (2015) shows that this specification is suitable to capturing interest-rate volatility without increasing the number of factors. The reason for not considering the affine specification (i.e., square-root of some of the elements of X_t) is addressed below. We label the model consisting of (1) and (11) as "NS-SV" (Nelson-Siegel Spanned-Volatility) and the model consisting of (7), (11), and (12) as "AF-SV" (Arbitrage-Free Spanned-Volatility). The NS/AF-CV model is a special case of the NS/AF-SV model with $\beta_i = 0$ for all i.

[3] We also consider a level-dependent correlation matrix satisfying the positive definiteness. The results (available upon request) show that the effect of no-arbitrage on the NS yield curve is not magnified by the extension to spanned volatility and correlation.

Since the AF-SV model is non-affine, there is no closed-form of the no-arbitrage yield, denoted as $Y^{AFSV}(X_t, \tau)$. We then compute this value by relying on an approximation method proposed and implemented by Takamizawa and Shoji (2009), and Takamizawa (2018). The method approximates conditional moments as a system of ordinary differential equations. Because the zero-coupon bond price is the conditional first moment of the stochastic discount factor, this method can directly be applied. Appendix A presents the accuracy of the approximation, which remains as long as we use realistic values for the factors and parameters.

In the spanned-volatility model, the covariance matrix $S_t R S_t$ is guaranteed to be positive definite for any value of X_t without restricting the sign of X_t. This is in sharp contrast to affine term structure models that require sign constraints, by which we cannot adopt the same risk-neutral drift as in (6). When using a different risk-neutral drift, it is difficult to distinguish between the volatility channel and the (risk-neutral) drift channel of no-arbitrage effects. Thus, this study does not employ the affine models with spanned volatility.

3 Data

This study uses the same dataset constructed by DL, which is downloaded from Francis Diebold's website. It consists of end-of-month zero-coupon bond yields with maturities of 1, 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months, and covers the period from January 1985 to December 2000 (192 observations).[4]

To check the robustness, we use an alternative dataset constructed by GSW (2007). [5] To maintain consistency with the main dataset, we use end-of-month observations of continuously compounded zero-coupon bond yields with maturities ranging from one to ten years, for the period from January 1985 to December 2017 (396 observations). Previous studies that limit the maturity range to ten years include Bauer, et al. (2012), Dai and Singleton (2000), Diebold et al. (2006), Diebold et al. (2008), Duffee (2002), Joslin et al. (2011), Koopman et al. (2010), and Wu and Xia (2016). The summary of the results is presented in Appendix B. Since the period of low volatility is longer than it is in the main dataset, the effect of no-arbitrage is smaller for both the constant- and spanned-volatility models. More important here is that it differs little between these models, confirming our conclusion that spanned volatility does not magnify the effect of no-arbitrage.

[4]While the data start from January 1970, DL use them from January 1985.

[5]The data are downloaded from https://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html.

4 Extraction of factors and estimation of their dynamics

We consider two approaches to extract the factors and estimate the parameters. Section 4.1 explains the two-step approach demonstrated by DL, in which the factors are first extracted by fitting the NS yield curve to the observed curve and then their dynamics are estimated using the data pooled from the first step. These factors and parameters from the NS-CV/SV models are fed into the AF-CV/SV models to compute the no-arbitrage yield curve, which is finally compared to the original NS yield curve. The two-step approach examines the effect of no-arbitrage from the perspective of a *conventional* use of the NS model.

Section 4.2 explains the one-step approach, in which we estimate the NS-CV/SV models (without no-arbitrage) and the AF-CV/SV models (with no-arbitrage) to obtain the values of the factors and parameters simultaneously. The model-implied yields would differ little with and without no-arbitrage because they match the observed values directly. Instead, the model-implied factors and parameters would differ. These differences are not fully investigated by previous studies, especially under spanned volatility. The one-step approach aims to find the effect of no-arbitrage from the perspective of a *sophisticated* use of the NS model.

4.1 Two-step approach

Let $y_{t,\tau}$ be a τ-maturity zero-coupon bond yield observed at time t. Then, in the first step, the vector X_t of factors is obtained at each point in time from the following regression equation:

$$y_{t,\tau} = B(\tau)' X_t + \epsilon_{t,\tau} \ , \tag{14}$$

where $B(\tau)$ is given by (2) and $\epsilon_{t,\tau}$ is a measurement error. Following DL, we fix the value of λ in $B(\tau)$ at $\lambda = 0.0609$ and implement the ordinary least squares (OLS) $T(= 192)$ times to obtain the T-by-3 data on X_t. We express X_t in raw numbers rather than in percentage terms.

In the second step, the dynamics of X_t, given by (3) for constant volatility and (11) for spanned volatility, are estimated using the data on X_t above. The (quasi-)maximum likelihood method is employed, in which the conditional first and second moments are substituted into the normal density function. While the likelihood function is exact for the NS-CV model, it is not for the NS-SV model. However, the moments are computed exactly as the drift vector is linear and the instantaneous covariance matrix is quadratic in X_t.

4.2 One-step approach

We also employ the (quasi-)maximum likelihood method together with the (extended) Kalman filter. The transition equation is based on (3) for constant volatility and (11) for spanned volatility. The observation equation is

$$y_{t,\tau} = Y^j(X_t, \tau) + \epsilon_{t,\tau} \ , \qquad \epsilon_{t,\tau} \sim N(0,\ \omega_\tau^2) \qquad (j = \{NS, AFCV, AFSV\}) \ . \tag{15}$$

Several notes on the estimation are in order. First, λ in K_1^Q is fixed at $\lambda = 0.0609$ as in DL, rather than treated as a free parameter in order to control for the risk-neutral drift. Second, to ease the interpretation of the results, the standard deviations of measurement errors are simplified to ω_1 for $\tau = 3$ (month) and ω_2 for the remaining maturities. The three-month yield is separated because it originally has a larger measurement error when explained by the NS model. Third, the initial value of X_t is given by the observed value, which is obtained by fitting the NS model to the first observation of the yield curve, rather than by the model-implied unconditional mean. Hence, the likelihood function is conditional on the first observation, which greatly enhances numerical stability. Finally, to apply the Kalman filter, the no-arbitrage yield for the spanned-volatility model, $Y^{AFSV}(X_t, \tau)$, which is a nonlinear function of X_t, is linearized around a predicted value of X_t. More sophisticated, computationally demanding estimation methods are difficult to apply when combined with the computation of $Y^{AFSV}(X_t, \tau)$.

5 Results

We report the results for the two approaches explained in Section 4. Section 5.1 compares the ex-post no-arbitrage yield to the original NS yield. We also address how the factor loadings change by extending the volatility from constant to spanned. Section 5.2 compares the factors and parameters with and without no-arbitrage obtained in the one-step approach.

5.1 Results for the two-step approach

Table 1 presents the estimates (standard errors) of the parameters for the two-step approach. In this study, the standard error is computed by the outer product of the gradient of the log-likelihood function. For parsimony, only the diagonal elements of K_1 (i.e., $k_{1,ii};\ i = 1, 2, 3$) are estimated. Furthermore, since $k_{0,2}$ and $k_{0,3}$ in K_0 are statistically insignificant when estimated as free parameters, only $k_{0,1}$ is estimated. NS-SV-F is the fully-parameterized version whereas

NS-SV-R is a restricted version, where only significant parameters in the volatility function $s(X_t)$ are estimated.

The results for the NS-CV model show that the curvature factor $x_{3,t}$ is the most volatile ($\sigma_3 = 0.026$) and least persistent ($k_{1,33} = -1.158$) among the factors. The level factor $x_{1,t}$ and the slope factor $x_{2,t}$ are similarly volatile ($\sigma_1 = \sigma_2 = 0.011$) while the latter is somewhat more persistent ($k_{1,11} = -0.308 < -0.163 = k_{1,22}$). Only the pair of $x_{1,t}$ and $x_{2,t}$ has a significant correlation coefficient ($\rho_{12} = -0.673$).

These basic features of the factors remain in the spanned-volatility model. In particular, the estimates in the drift do not differ significantly between the constant- and spanned-volatility models. It is the overall statistical fit that improves owing to the extension to spanned volatility. The value of the maximum log-likelihood (LogL) for the NS-SV-R model is 2411, increased from that for the NS-CV model (2391) without increasing the number of free parameters. However, more volatility parameters contribute little to the overall fit because the LogL value for the NS-SV-F model is larger than that for the NS-SV-R model by only 2.

Among the parameters in $s_i(X_t)$, only β_{11} and β_{21} are significant. For the NS-SV-R model, these values are 0.139 and 0.149, respectively. This result implies that the level factor alone is a significant driver for the volatility of changes in the level and slope factors in this dataset. These factors are more volatile for higher levels of interest rates. By contrast, such a level dependence of volatility is not observed for the curvature factor because none of β_{3i} ($i = 1, 2, 3$) are significant.

These estimates in Table 1, together with the data on X_t, are used as inputs for the AF-CV/SV models to obtain the ex-post no-arbitrage yield. Table 2 presents the mean, standard deviation (S.D.), minimum (Min), and maximum (Max) of the difference in the model-implied yields, $Y^j(X_t, \tau) - Y^{NS}(X_t, \tau)$ ($j = \{AFCV, AFSV\}$). Remember that the difference is constant for the AF-CV model at a given maturity (i.e., $A(\tau)$). For the AF-CV model, the difference is negative and more so for longer maturity by construction of $A(\tau)$ given by (10). It ranges from almost zero at $\tau = 3$ (months) to -20 basis points (bps, 1 bp = 1/10,000) at $\tau = 120$. This gives a benchmark for the extent to which a conventional use of the NS model is inconsistent with no-arbitrage when using the DL dataset. Coroneo et al. (2011) conclude that the inconsistency of this magnitude may not be a serious concern.

The mean difference for the AF-SV models has a similar pattern. For instance, the average ten-year yield generated by these models is lower by around -20 bps than that for the NS model.

Unlike the constant-volatility model, the spanned-volatility model generates a varying difference because it has the factor loadings that are no longer the same as those for the original NS model, despite having the same risk-neutral drift. The standard deviation of the difference is 6.6 bps for AF-SV-F and 4.5 bps for AF-SV-R at $\tau = 120$. While the smallest deviation (Max in Table 2) is similar between the two models (e.g., around -9.5 bps at $\tau = 120$), the largest deviation (Min in Table 2) for the unrestricted model is larger (-41 bps) than that for the restricted model (-32 bps). Such a large deviation is observed in the early part of the sample, when the level of interest rates is high.

Figure 1 depicts the time-series of the difference in ten-year yields between the spanned-volatility and NS models. For reference, the analogous difference generated by the AF-CV model is also shown, which is the horizontal line at around -20 bps. Apparently, the difference with and without no-arbitrage trends upward, which aligns with a decreasing trend in the level of interest rates. The proportion of cases in which the difference for the spanned-volatility model is smaller in absolute value than that for the constant-volatility model (that is, the frequency at which the plot is inside the horizontal line) is 48% for AF-SV-F and 66% for AF-SV-R. Therefore, the extension to spanned volatility does not necessarily reinforce the effect of no-arbitrage on the NS yield curve.

To explore why the interest-rate level matters with the deviation in yields between the AF-SV and NS models, we project $Y^{AFSV}(X_t, \tau)$ on a constant and X_t for each τ:

$$Y^{AFSV}(X_t, \tau) = c_{0,\tau} + c_\tau' X_t + u_{t,\tau} \ , \tag{16}$$

where the residual term $u_{t,\tau}$ collects the remaining nonlinear terms of X_t. The coefficients are estimated with OLS.

Figure 2 depicts the estimated $c_{0,\tau}$ and c_τ over τ. For comparison, the analogous plots for the NS and AF-CV models are displayed: $c_{0,\tau} = 0$ and $c_{1,\tau} = B(\tau)$ for NS, and $c_{0,\tau} = A(\tau)$ and $c_{1,\tau} = B(\tau)$ for AF-CV, where $A(\tau)$ and $B(\tau)$ are given by (10) and (2), respectively. Panel (a) presents $c_{0,\tau}$ expressed in bps. Notice that for the AF-CV model, it is equal to the difference presented in Table 1. It is found that $c_{0,\tau}$ is positive for the spanned-volatility model and more so for longer maturities. For instance, at $\tau = 120$, it is 12 bps for AF-SV-F and 5 bps for AF-SV-R. The positive constant, however, is offset by a smaller $c_{1,\tau}$ (the coefficient of $x_{1,t}$) in Panel (b). At $\tau = 120$, it is 0.95 for AF-SV-F and 0.97 for AF-SV-R, both of which are smaller than that for AF-CV fixed at one. Therefore, when $x_{1,t}$ is large (small), the yield for the spanned-volatility model is low (high) relative to that for the constant-volatility model. The reason for the decreasing $c_{1,\tau}$ is intuitively understood by recalling the affine term structure model with spanned volatility. In estimating this model, the resulting volatility factor is often persistent and correlated with long-term interest rates. In short, a factor driving the volatility is also the level factor of interest rates. Then, the convexity adjustment matters with the loading on the level factor as well as the constant term in the no-arbitrage yield function, and the loading decreases with maturity as does the constant term.

By contrast, $c_{2,\tau}$ and $c_{3,\tau}$ in Panels (c) and (d) differ little across the models. This holds true for the AF-SV-F model, in which the slope and curvature factors also drive the volatility, thereby their loadings may possibly be affected by no-arbitrage as the convexity adjustment.

In summary, the ex-post no-arbitrage yield is no longer the same as the original yield computed with the NS model. However, how different they are differs little on average for constant and spanned volatility.

5.2 Results for the one-step approach

5.2.1 Parameter estimates for the constant-volatility model

Table 3 presents the estimates (standard errors) of the parameters in the one-step approach, in which the filtration of X_t from the cross-section of interest rates and estimation of the dynamics of X_t is performed simultaneously. First, we compare the results for the NS-CV model between Tables 1 and 3. The estimates do not change significantly, indicating that the dynamics of X_t are not much affected by how X_t is elicited, simultaneously with or independently of the estimation of the dynamics.

Next, we compare the results between NS-CV and AF-CV within Table 3, and find some notable changes. First, the estimate of σ_1 decreases from 0.010 without no-arbitrage to 0.008 with it. Since a lower volatility leads to a smaller convexity adjustment given the factor loading $B(\tau)$ (i.e., $\lambda = 0.0609$ in (2)), the no-arbitrage condition works to reduce the convexity adjustment by affecting the volatility of the level factor. Put differently, the volatility estimated from the time-series dimensions of the data alone is large for the cross-sectional dimensions of the data that are consistent with no-arbitrage.

Second, the value of the maximum log-likelihood (LogL) decreases from 18037 to 17933 by imposing no-arbitrage, indicating that no-arbitrage is restrictive for the overall statistical fit.[6]

[6]This result is not conclusive because in the alternative dataset, which contains more observations of low interest rates, the imposition of no-arbitrage does not necessarily reduce the LogL value, which is explained as

Meanwhile, the magnitude of the measurement errors does not appear to change largely. ω_1 for the three-month yield and ω_2 for the remaining yields increase by 0.4 bps and 0.2 bps, respectively, by imposing no-arbitrage. Though economically small, the latter increase is statistically significant, reducing the LogL value by about 100.

5.2.2 Parameter estimates for the spanned-volatility model

It is easy to see that the estimates for the NS-SV-R model do not change significantly between Tables 3 and 1. This also holds true for the NS-SV-F model, although it is not straightforward to see this because of the many insignificant parameters. Therefore, the finding that the estimated dynamics of X_t are robust to different eliciting approaches also holds for spanned volatility.

We next compare the results between NS-SV-R and AF-SV-R in the last two columns of Table 3. As is the case for the constant-volatility model, the imposition of no-arbitrage mainly affects the volatility of changes in the level factor, which here is a function of X_t or $s_1(X_t)$ given in (13). The estimate of β_{11} decreases from 0.133 to 0.097 by imposing no-arbitrage. Given the level of $x_{1,t}$ (around 0.077), the decrease in volatility is of similar magnitude to that for the constant-volatility model. Furthermore, while the increase in measurement errors is marginal, this decreases the LogL value by about 100.

The results for the unrestricted spanned-volatility model in the middle two columns of Table 3 have a similar pattern to the restricted counterpart. The volatility of the level factor decreases by imposing no-arbitrage. More precisely, though the constant term becomes positive and marginally significant with no-arbitrage ($\sigma_1 = 0.002$), the coefficient of the level factor is much smaller with it ($\beta_{11} = 0.057$) than without it ($\beta_{11} = 0.195$). Given the level of $x_{1,t}$, the decrease in β_{11} dominates the increase in σ_1. The unrestricted model has other parameters that become significant by imposing no-arbitrage. β_{12} decreases from 0.074 to -0.089. Since the slope factor is $x_{2,t} = r_t - x_{1,t}$ (i.e., the spread between short- and long-term yields), the negative estimate of β_{12} implies that a negative shock to $x_{2,t}$—a shock that raises the slope of the yield curve—raises the volatility of the level factor. Although the level of β_{33} does not change (-0.19), the estimation precision improves with no-arbitrage. Because some of the originally insignificant parameters are used for fitting the data, no-arbitrage may be less restrictive for AF-SV-F. Indeed, the decrease

follows. The low interest rate environment reduces the significance of the parameters in the covariance matrix. However, when no-arbitrage is imposed, some of these parameters become significant as they are adjusted to the cross-section of interest rates rather than the time-series. Then, the improvement in the cross-sectional fit, though economically small, increases the LogL value.

in LogL is 89 (from 18060 to 17971), which is a bit less than that for AF-CV and AF-SV-R.

Taken together, while no-arbitrage may be restrictive from a statistical point of view, the way and extent to which no-arbitrage affects the time-series and cross-section of interest rates is similar for both the constant- and spanned-volatility models.

5.2.3 Elicited factors

The difference in factors with and without no-arbitrage is denoted as $e^j_{i,t} = x^{AFj}_{i,t} - x^{NSj}_{i,t}$ ($i = 1, 2, 3$; $j = \{CV, SVF, SVR\}$). Table 4 presents the mean, standard deviation (S.D.), minimum (Min), maximum (Max), and first autocorrelation (AR(1)) of $e^j_{i,t}$, where the numbers are in bps except for the AR(1). For each factor, the mean difference is similar across the models. For instance, for the level factor, it is 13 bps, which is less than the difference for the ten-year yield by the two-step approach (around 20 bps). The mean difference for the curvature factor is the largest in absolute value, implying that it is difficult to pin down precisely. However, since the loading on the curvature factor is small, such a large difference does not lead to a large difference in yields.

Unlike the mean, the standard deviation of the difference in factors varies between the constant- and spanned-volatility models. The standard deviation of $e^{SVR}_{1,t}$ is 3 bps, which is larger than that for the constant-volatility model (0.8 bps). Also of note is that the spanned-volatility model generates a much more persistent difference than does the constant-volatility model. The first autocorrelation of $e^{SVR}_{i,t}$ ranges from 0.85 ($i = 3$) to 0.98 ($i = 2$), whereas that of $e^{CV}_{i,t}$ ranges from 0.15 ($i = 2$) to 0.46 ($i = 3$).

Figure 3 depicts the time-series of $e^j_{i,t}$. $e^{CV}_{i,t}$ slightly fluctuates around the mean, whereas $e^{SV}_{i,t}$ trends toward zero, which is in line with Figure 1 using the two-step approach. These results also support the finding that spanned volatility does not necessarily magnify the effect of no-arbitrage on the filtration of factors.

6 Concluding remarks

This study examined the effect of no-arbitrage on the NS (1987) yield curve when the volatility of changes in interest-rate factors depends on the level of the factors; that is, spanned volatility. Our primary focus is whether spanned volatility magnifies the no-arbitrage effect relative to constant volatility. The question is important to address because if it were *not* the case, it could be possible to enrich the dynamics of interest rates based on the NS model, such as by augmenting stochastic volatility for time-series after extracting the factors from the cross-section. Although such an augmented model might not escape the inconsistency with no-arbitrage, the degree of inconsistency would not be very different from that for a constant-volatility model used in the conventional approach demonstrated by DL (2006).

We find evidence supporting this argument by using two approaches with the DL dataset. In the two-step approach, in which we extract the factors and estimate their dynamics separately without imposing no-arbitrage, the ex-post no-arbitrage yield deviates from the original NS yield by 20 bps on average at the ten-year maturity for both the constant- and spanned-volatility models. The deviation for the latter model varies over time and becomes occasionally large. However, the frequency at which the deviation exceeds 20 bps is around or less than 50 percent. In the one-step approach, in which we extract the factors and estimate their dynamics simultaneously, the volatility of changes in the level factor declines by imposing no-arbitrage, which is of similar magnitude on average for the constant- and spanned-volatility models. Furthermore, while the extracted factors may differ with and without no-arbitrage, the overall difference is not magnified by extending from constant to spanned volatility.

Appendix A. Accuracy of approximation

This appendix examines the accuracy of the approximation used to compute the arbitrage-free yield for AF-SV-F and AF-SV-R. The parameter values are taken from Table 3. An approximation error is defined as the difference between the approximate yield and the yield computed by the Monte Carlo (MC) method.

To compute the MC yield, a path of the state vector $\{X_s\}_t^{t+\tau}$ is generated from the risk-neutral distribution given by (12), where dt is replaced by $\Delta t = 1/240$, which is roughly equal to a daily observation frequency. Several starting values of $X_t = (x_{1,t},\ x_{2,t},\ x_{3,t})$ are selected as follows. First, we pick the three sets of X_t containing the minimum, median, and maximum values of the level factor $x_{1,t}$. This is repeated for the slope factor $x_{2,t}$ and the curvature factor $x_{3,t}$, resulting in nine sets of X_t. The data on X_t are obtained by fitting the NS yield curve to the observed curve at each point in time over the sample period. The number of repetition is set at 100,000 with antithetic variates.

Table A1 presents the ten-year yields computed using the approximation and MC methods, and the difference between the two as an approximation error. We focus on only the ten-year maturity as the approximation worsens as the maturity period increases. The accuracy of the approximation tends to be low when either the slope or curvature factor is low and when the level factor is high; the difference then exceeds 2 bps. Note that since the slope factor is $x_{2,t} = r_t - x_{1,t}$ given by (7), the minimum (maximum) slope corresponds to the steepest (least steep) yield curve. On typical days when one of the factors takes the median value, the difference is less than 1 bp. Overall, accuracy does not seem to be a serious concern.

Appendix B. Robustness check using an alternative dataset

The dataset constructed by GSW (2007) are used to check the robustness of the results. We use end-of-month observations of continuously compounded zero-coupon bond yields with maturities ranging from one to ten years for the period from January 1985 to December 2017 (396 observations). Compared to DL's dataset, this dataset contains more observations of low interest rates.

We modify the estimation slightly. In the one-step approach with the (extended) Kalman filter, the observation equation without no-arbitrage is

$$y_{t,\tau} = c_0 + B(\tau)' X_t + \epsilon_{t,\tau} \ . \tag{17}$$

That is, we add a constant c_0. Similarly, under no-arbitrage, the risk-neutral distribution of dX_t is

$$dX_t \sim N[\ (K_0^Q + K_1^Q X_t)\ dt, \quad S_t R S_t\ dt\ ]\ , \tag{18}$$

where $K_0^Q = (k_{0,1}^Q,\ 0,\ 0)'$. That is, we add $k_{0,1}^Q$. These additional parameters associated with the cross-section of interest rates are highly significant, which is shown in Table B3, and greatly improve the statistical fit to the alternative data.

Tables B1–B4 and Figures B1–B3 are analogous to Tables 1–4 and Figures 1–3 of the main text, respectively. We do not present the results for the unconstrained spanned-volatility model (NS/AF-SV-F) in the one-step approach as they are similar to those for the restricted counterpart (NS/AF-SV-R). In summary, while no-arbitrage more or less affects the yield curve and its dynamics, the extent to which it does is similar between constant and spanned volatility, which is the same conclusion drawn from the main dataset.

References

Bauer, M. D., G. D. Rudebusch, and J. C. Wu, 2012, Correcting Estimation Bias in Dynamic Term Structure Models, *Journal of Business & Economic Statistics* 30, 454-467.

Björk, T., and B. J. Christensen, 1999, Interest rate dynamics and consistent forward rate curves, *Mathematical Finance* 9, 323-348.

Chen, R., and K. Du, 2013, A generalized arbitrage-free Nelson-Siegel model: The impact of unspanned stochastic volatility, *Finance Research Letters* 10, 41-48.

Christensen, J. H., F. X. Diebold, and G. D. Rudebusch, 2009, An arbitrage-free generalized Nelson-Siegel term structure model, *Econometrics Journal* 12, C33-C64.

Christensen, J. H., F. X. Diebold, and G. D. Rudebusch, 2011, The affine arbitrage-free class of Nelson-Siegel term structure models, *Journal of Econometrics* 164, 1-20.

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

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

Dai, Q., and K. J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943-1978.

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

Diebold, F. X., C. Li, and V. Z. Yue, 2008, Global yield curve dynamics and interactions: A dynamic Nelson-Siegel approach, *Journal of Econometrics* 146, 351-363.

Diebold, F. X., G. D. Rudebusch, and S. B. Aruoba, 2006, The macroeconomy and the yield curve: A dynamic latent factor approach, *Journal of Econometrics* 131, 309-338.

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

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

Gürkaynak, R. S., B. Sack, and J. H. Wright, 2007, The U.S. treasury yield curve: 1961 to the present, *Journal of Monetary Economics* 54, 2291-2304.

Heath, D., R. Jarrow, and A. Morton, 1992, Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation, *Econometrica* 60, 77-105.

Joslin, S., K. J. Singleton, and H. Zhu, 2011, A New Perspective on Gaussian Dynamic Term Structure Models, *Review of Financial Studies* 24, 926-970.

Koopman, S. J., M. I. P. Mallee, and M. Van der Wel, 2010, Analyzing the term structure of interest rates using the dynamic Nelson-Siegel model with time-varying parameters, *Journal of Business & Economic Statistics* 28, 329-343.

Nelson, C. R., and A. F. Siegel, 1987, Parsimonious modeling of yield curves, *Journal of Business* 60, 473-489.

Söderlind, P., and L. Svensson, 1997, New techniques to extract market expectations from financial instruments, *Journal of Monetary Economics* 40, 383-429.

Takamizawa, H., 2015, Predicting Interest Rate Volatility Using Information on the Yield Curve, *International Review of Finance* 15, 347-386.

Takamizawa, H., 2018, A Term Structure Model of Interest Rates with Quadratic Volatility, *Quantitative Finance* 18, 11731198.

Takamizawa, H., and I. Shoji, 2009, Modeling the term structure of interest rates with general diffusion processes: A moment approximation approach, *Journal of Economic Dynamics and Control* 33, 65-77.

Wu, J. C., and F. D. Xia, 2016, Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound, *Journal of Money, Credit and Banking* 48, 253-291.

	NS-CV		NS-SV-F		NS-SV-R	
$k_{1,11}$	−0.308	(0.122)	−0.231	(0.131)	−0.241	(0.121)
$k_{1,22}$	−0.163	(0.100)	−0.143	(0.110)	−0.174	(0.114)
$k_{1,33}$	−1.158	(0.344)	−1.100	(0.348)	−1.129	(0.345)
$k_{0,1}$	0.019	(0.010)	0.014	(0.009)	0.014	(0.008)
$\sigma_1 \times 10^2$	1.064	(0.044)	−0.208	(0.224)		
$\sigma_2 \times 10^2$	1.126	(0.059)	−0.158	(0.348)		
$\sigma_3 \times 10^2$	2.610	(0.115)	1.924	(1.204)	2.607	(0.115)
ρ_{12}	−0.673	(0.029)	−0.718	(0.033)	−0.713	(0.029)
ρ_{13}	0.104	(0.084)	0.120	(0.082)	0.118	(0.078)
ρ_{23}	−0.076	(0.066)	−0.108	(0.064)	−0.100	(0.064)
β_{11}			0.188	(0.044)	0.139	(0.006)
β_{12}			0.084	(0.057)		
β_{13}			−0.050	(0.043)		
β_{21}			0.184	(0.059)	0.149	(0.009)
β_{22}			0.048	(0.061)		
β_{23}			0.010	(0.039)		
β_{31}			0.114	(0.162)		
β_{32}			0.091	(0.092)		
β_{33}			−0.041	(0.062)		
LogL	2391		2413		2411	

Table 1: Parameter estimates (standard errors) by the two-step approach without no-arbitrage
In the first step, the vector of interest-rate factors X_t is obtained by fitting the NS yield curve to the observed curve at each point in time. In the second step, the dynamics of X_t are estimated using the (quasi-)maximum likelihood method with the standard error computed by the outer product of the gradient of the log-likelihood function. The distribution of dX_t is given by $dX_t \sim N[(K_0 + K_1X_t)dt,\ S_tRS_t\ dt]$: only the first element of K_0 ($k_{0,1}$) and the diagonal elements of K_1 ($k_{1,ii}$) are estimated; R is a correlation matrix with the ij-th element ρ_{ij}; S_t is a diagonal matrix with the i-th diagonal element specified as $s(X_t) = \sigma_i + \beta_i' X_t$. NS-SV-F is a fully parameterized model, and NS-CV and NS-SV-R are models with restrictions on βs. The sample period is from 1985/1 to 2000/12 (192 observations).

τ	AF-CV	AF-SV-F				AF-SV-R			
mon.		Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	−0.1	0.0	0.0	0.0	−0.1	0.0
9	−0.1	−0.1	0.0	−0.2	0.0	−0.1	0.0	−0.2	0.0
12	−0.2	−0.2	0.1	−0.3	−0.1	−0.1	0.0	−0.3	−0.1
15	−0.3	−0.3	0.1	−0.5	−0.1	−0.2	0.1	−0.5	−0.1
18	−0.4	−0.4	0.1	−0.8	−0.2	−0.4	0.1	−0.7	−0.2
21	−0.6	−0.6	0.2	−1.2	−0.3	−0.6	0.1	−1.0	−0.3
24	−0.8	−0.8	0.2	−1.6	−0.4	−0.8	0.2	−1.3	−0.5
30	−1.3	−1.4	0.4	−2.7	−0.8	−1.3	0.3	−2.2	−0.8
36	−2.0	−2.1	0.6	−4.2	−1.1	−1.9	0.4	−3.3	−1.2
48	−3.6	−3.9	1.2	−8.0	−2.1	−3.5	0.7	−6.1	−2.2
60	−5.6	−6.2	1.9	−12.9	−3.1	−5.4	1.2	−9.6	−3.3
72	−7.9	−8.8	2.8	−18.6	−4.2	−7.6	1.8	−13.8	−4.5
84	−10.4	−11.7	3.8	−24.8	−5.4	−10.0	2.4	−18.3	−5.7
96	−13.3	−14.8	4.9	−31.1	−6.6	−12.6	3.1	−23.0	−6.9
108	−16.4	−18.1	5.8	−36.7	−7.9	−15.2	3.8	−27.7	−8.2
120	−19.9	−21.3	6.6	−40.9	−9.3	−18.0	4.5	−31.8	−9.6

Table 2: Difference between arbitrage-free and NS yields
Mean, standard deviation (S.D.), minimum (Min), and maximum (Max) of the difference in the model-implied yields with and without no-arbitrage are presented in bps. The values of the interest-rate factors X_t and parameters in Table 1 are fed into the no-arbitrage version of the models to compute the no-arbitrage yield, which is then compared to the original NS yield. For AF-CV, the difference is constant at a given maturity (i.e., $A(\tau)$ in (10)). The sample period is from 1985/1 to 2000/12 (192 observations).

	NS-CV		AF-CV		NS-SV-F		AF-SV-F		NS-SV-R		AF-SV-R	
$k_{1,11}$	−0.260	(0.114)	−0.228	(0.070)	−0.168	(0.120)	−0.162	(0.072)	−0.168	(0.115)	−0.156	(0.068)
$k_{1,22}$	−0.164	(0.100)	−0.168	(0.091)	−0.164	(0.105)	−0.165	(0.084)	−0.170	(0.113)	−0.134	(0.100)
$k_{1,33}$	−1.110	(0.320)	−1.003	(0.293)	−1.227	(0.381)	−1.224	(0.335)	−1.121	(0.317)	−1.101	(0.281)
$k_{0,1}$	0.016	(0.010)	0.014	(0.006)	0.010	(0.008)	0.009	(0.005)	0.009	(0.008)	0.009	(0.005)
$\sigma_1 \times 10^2$	1.031	(0.047)	0.763	(0.019)	−0.317	(0.250)	0.171	(0.089)				
$\sigma_2 \times 10^2$	1.119	(0.064)	1.053	(0.060)	−0.139	(0.340)	−0.228	(0.323)				
$\sigma_3 \times 10^2$	2.414	(0.146)	2.369	(0.146)	1.651	(1.222)	1.387	(0.798)	2.408	(0.145)	2.265	(0.130)
ρ_{12}	−0.673	(0.034)	−0.607	(0.031)	−0.717	(0.035)	−0.676	(0.030)	−0.709	(0.033)	−0.627	(0.031)
ρ_{13}	0.196	(0.084)	−0.099	(0.064)	0.230	(0.090)	0.178	(0.058)	0.222	(0.078)	0.059	(0.053)
ρ_{23}	−0.081	(0.073)	0.130	(0.073)	−0.116	(0.079)	−0.106	(0.070)	−0.113	(0.072)	−0.045	(0.069)
β_{11}					0.195	(0.048)	0.057	(0.015)	0.133	(0.006)	0.097	(0.002)
β_{12}					0.074	(0.061)	−0.089	(0.025)				
β_{13}					−0.048	(0.047)	0.002	(0.025)				
β_{21}					0.185	(0.062)	0.188	(0.058)	0.148	(0.009)	0.133	(0.008)
β_{22}					0.070	(0.066)	0.075	(0.057)				
β_{23}					−0.015	(0.043)	−0.024	(0.035)				
β_{31}					0.129	(0.175)	0.119	(0.110)				
β_{32}					0.127	(0.131)	0.086	(0.085)				
β_{33}					−0.189	(0.110)	−0.187	(0.085)				
$\omega_1 \times 10^4$	14.7	(0.680)	15.1	(0.739)	14.8	(0.727)	15.1	(0.807)	14.7	(0.669)	15.0	(0.718)
$\omega_2 \times 10^4$	6.6	(0.058)	6.8	(0.066)	6.6	(0.059)	6.8	(0.065)	6.6	(0.054)	6.8	(0.061)
LogL	18037		17933		18060		17971		18056		17959	

Table 3: Parameter estimates (standard errors) by the one-step approach with (AF-) and without (NS-) no-arbitrage
The (quasi-)maximum likelihood method with the (extended) Kalman filter is used for estimating the parameters with the standard error computed by the outer product of the gradient of the log-likelihood function. The observation equation is given in (15) with the standard deviations of measurement errors given by ω_1 for $\tau = 3$ (month) and ω_2 for the remaining maturities. The transition equation is from $dX_t \sim N[(K_0 + K_1 X_t)dt,\ S_t R S_t\ dt]$: only the first element of K_0 ($k_{0,1}$) and the diagonal elements of K_1 ($k_{1,ii}$) are estimated; R is a correlation matrix with the ij-th element ρ_{ij}; S_t is a diagonal matrix with the i-th diagonal element specified as $s(X_t) = \sigma_i + \beta_i' X_t$. NS/AF-SV-F is a fully parameterized model, and NS/AF-CV and NS/AF-SV-R are models with restrictions on βs. The sample period is from 1985/1 to 2000/12 (192 observations).

	e_1			e_2			e_3		
	CV	SV-F	SV-R	CV	SV-F	SV-R	CV	SV-F	SV-R
Mean	13.0	13.0	13.1	−10.6	−10.8	−10.6	−25.0	−24.1	−25.0
S.D.	0.8	3.4	3.0	0.6	2.8	2.3	2.3	7.8	6.4
Min	9.3	4.2	4.5	−12.1	−17.0	−18.4	−32.1	−53.4	−42.0
Max	15.6	20.6	22.2	−8.9	−5.0	−5.4	−13.0	1.0	−2.8
AR(1)	0.32	0.89	0.94	0.15	0.98	0.98	0.46	0.68	0.85

Table 4: Difference in factors with and without no-arbitrage
Mean, standard deviation (S.D.), minimum (Min), maximum (Max), and first auto-correlation (AR(1)) of the difference in the model-implied factors with and without no-arbitrage are presented: the numbers are in bps except for AR(1). The difference is computed for each factor i as $e_{i,t}^{j} = x_{i,t}^{AFj} - x_{i,t}^{NSj}$ $(i = 1, 2, 3;\ j = \{CV, SVF, SVR\})$. CV stands for a model with constant volatility, and SV-F and SV-R stand for models with spanned volatility with full and restricted parameters, respectively. The sample period is from 1985/1 to 2000/12 (192 observations).

State	AF-SV-F Approx (%)	MC (%)	diff (bps)	AF-SV-R Approx (%)	MC (%)	diff (bps)
Min Level	4.339	4.339	0.0	4.344	4.344	0.0
Min Slope	7.756	7.734	2.2	7.767	7.745	2.2
Min Curvature	6.584	6.562	2.3	6.624	6.604	2.0
Med Level	7.137	7.134	0.3	7.127	7.123	0.5
Med Slope	8.742	8.737	0.4	8.706	8.699	0.8
Med Curvature	6.747	6.746	0.0	6.730	6.728	0.2
Max Level	11.633	11.617	1.6	11.562	11.529	3.3
Max Slope	5.356	5.356	0.0	5.351	5.350	0.0
Max Curvature	7.536	7.535	0.0	7.503	7.501	0.2

Table A1: Approximation error for the ten-year yield
The ten-year yields are computed using an approximation (Approx) method and the Monte Carlo (MC) method, respectively, for AF-SV-F and AF-SV-R with the parameter values given in Table 3. The difference between the two yields is expressed in bps. The yields are evaluated at nine sets of factors $X_t = (x_{1,t},\ x_{2,t},\ x_{3,t})$: for instance, "Min Level" represents a set in which $x_{1,t}$ takes the minimum, "Med Slope" represents a set in which $x_{2,t}$ takes the median, and "Max Curvature" represents a set in which $x_{3,t}$ takes the maximum. The data on X_t are obtained by fitting the NS yield curve to the observed curve at each point in time over the period from 1985/1 to 2000/12.

	NS-CV		NS-SV-F		NS-SV-R	
$k_{1,11}$	−0.127	(0.065)	−0.141	(0.057)	−0.148	(0.053)
$k_{1,22}$	−0.109	(0.074)	−0.137	(0.079)	−0.145	(0.074)
$k_{1,33}$	−0.138	(0.118)	−0.025	(0.154)	−0.019	(0.152)
$k_{0,1}$	0.004	(0.005)	0.006	(0.004)	0.006	(0.003)
$\sigma_1 \times 10^2$	1.133	(0.029)	0.099	(0.068)		
$\sigma_2 \times 10^2$	1.341	(0.040)	0.226	(0.152)		
$\sigma_3 \times 10^2$	3.013	(0.077)	0.064	(0.363)		
ρ_{12}	−0.700	(0.023)	−0.729	(0.025)	−0.735	(0.020)
ρ_{13}	−0.176	(0.040)	−0.081	(0.052)	−0.083	(0.048)
ρ_{23}	−0.093	(0.046)	−0.149	(0.053)	−0.147	(0.050)
β_{11}			0.126	(0.013)	0.137	(0.007)
β_{12}			0.014	(0.025)		
β_{13}			−0.081	(0.011)	−0.085	(0.008)
β_{21}			0.135	(0.024)	0.165	(0.007)
β_{22}			0.009	(0.019)		
β_{23}			−0.097	(0.015)	−0.115	(0.011)
β_{31}			0.317	(0.056)	0.319	(0.019)
β_{32}			0.051	(0.052)		
β_{33}			−0.305	(0.041)	−0.288	(0.031)
LogL	4811.9		4897.1		4894.7	

Table B1: Parameter estimates (standard errors) by the two-step approach without no-arbitrage for GSW data

This table is analogous to Table 1 of the main text, obtained with the dataset constructed by GSW (2007). The data are monthly, covering the period from 1985/1 to 2017/12 (396 observations).

τ	AF-CV	AF-SV-F				AF-SV-R			
year	Mean	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
1	−0.1	−0.1	0.0	−0.3	0.0	−0.1	0.1	−0.3	0.0
2	−0.6	−0.5	0.2	−1.6	−0.1	−0.5	0.3	−1.6	−0.1
3	−1.5	−1.3	0.6	−3.6	−0.2	−1.3	0.6	−3.8	−0.2
4	−2.7	−2.3	1.1	−5.9	−0.4	−2.3	1.1	−6.2	−0.3
5	−4.3	−3.5	1.7	−8.5	−0.5	−3.6	1.8	−8.9	−0.5
6	−6.2	−4.9	2.4	−12.2	−0.8	−5.0	2.5	−12.6	−0.6
7	−8.4	−6.5	3.1	−16.3	−1.0	−6.6	3.3	−16.7	−0.8
8	−10.9	−8.2	4.0	−20.6	−1.3	−8.3	4.1	−21.1	−1.1
9	−13.7	−10.1	4.8	−24.9	−1.6	−10.1	5.0	−25.4	−1.3
10	−16.9	−12.0	5.6	−28.8	−2.0	−12.0	5.9	−29.2	−1.6

Table B2: Difference between arbitrage-free and NS yields for GSW data
This table is analogous to Table 2 of the main text, obtained with the dataset constructed by GSW (2007). The data are monthly, covering the period from 1985/1 to 2017/12 (396 observations).

	NS-CV		AF-CV		NS-SV-R		AF-SV-R	
$k_{1,11}$	−0.130	(0.065)			−0.140	(0.059)	−0.142	(0.059)
$k_{1,22}$	−0.130	(0.085)			−0.129	(0.089)	−0.129	(0.090)
$k_{1,33}$	−0.300	(0.165)			−0.333	(0.238)	−0.333	(0.283)
$k_{0,1}$	0.003	(0.004)			0.005	(0.003)	0.004	(0.003)
$c_0 \times 10^2$	0.074	(0.001)			0.074	(0.001)		
$k_{0,1}^Q \times 10^2$			0.227	(0.005)			0.218	(0.003)
$\sigma_1 \times 10^2$	1.110	(0.029)	1.187	(0.034)	0.407	(0.055)	0.523	(0.039)
$\sigma_2 \times 10^2$	1.350	(0.042)	1.399	(0.046)	0.512	(0.124)	0.666	(0.107)
$\sigma_3 \times 10^2$	2.914	(0.079)	2.911	(0.076)	0.649	(0.302)	1.313	(0.277)
ρ_{12}	−0.716	(0.022)	−0.739	(0.022)	−0.734	(0.025)	−0.748	(0.021)
ρ_{13}	−0.142	(0.041)	−0.176	(0.042)	−0.067	(0.055)	−0.103	(0.047)
ρ_{23}	−0.071	(0.048)	−0.033	(0.049)	−0.097	(0.055)	−0.099	(0.056)
β_{11}					0.110	(0.011)	0.124	(0.005)
β_{12}								
β_{13}					−0.082	(0.011)	−0.021	(0.006)
β_{21}					0.136	(0.025)	0.131	(0.024)
β_{22}								
β_{23}					−0.102	(0.016)	−0.053	(0.014)
β_{31}					0.342	(0.060)	0.268	(0.055)
β_{32}								
β_{33}					−0.310	(0.043)	−0.471	(0.042)
$\omega \times 10^4$	2.98	(0.017)	2.95	(0.017)	2.97	(0.018)	2.78	(0.018)
LogL	23914		23937		23995		24153	

Table B3: Parameter estimates (standard errors) by the one-step approach with (AF-) and without (NS-) no-arbitrage for GSW data

This table is analogous to Table 3 of the main text, obtained with the dataset constructed by GSW (2007). The results for the fully-parameterized model (NS/AF-SV-F) are not presented as they are similar to those for NS/AF-SV-R presented above. The data are monthly, covering the period from 1985/1 to 2017/12 (396 observations).

	x_1		x_2		x_3	
	CV	SV-R	CV	SV-R	CV	SV-R
Mean	−28.8	−20.5	26.7	20.1	26.2	12.7
S.D.	0.1	10.9	0.1	5.8	0.5	29.1
Min	−29.2	−41.2	26.4	3.1	23.8	−62.4
Max	−28.3	9.6	26.9	30.5	28.4	69.6
AR(1)	0.05	0.98	0.26	0.97	0.07	0.99

Table B4: Difference in factors with and without no-arbitrage for GSW data
This table is analogous to Table 4 of the main text, obtained with the dataset constructed by GSW (2007). The results for the fully-parameterized model (NS/AF-SV-F) are not presented as they are similar to those for NS/AF-SV-R presented above. The data are monthly, covering the period from 1985/1 to 2017/12 (396 observations).

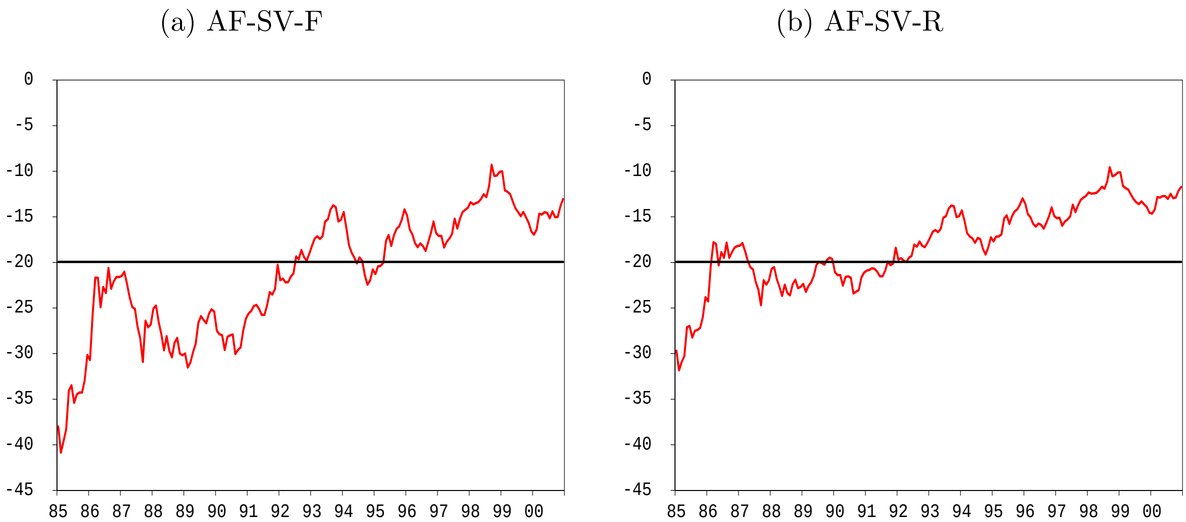

Figure 1: Time-series of difference in ten-year yields between spanned-volatility and NS models

The difference in ten-year yields is computed as $Y^{AFSV}(X_t, 120) - Y^{NS}(X_t, 120)$, where X_t is extracted by the two-step approach and the parameter values are given in Table 1 with $\lambda = 0.0609$. The sample period is from 1985/1 to 2000/12 (192 observations).

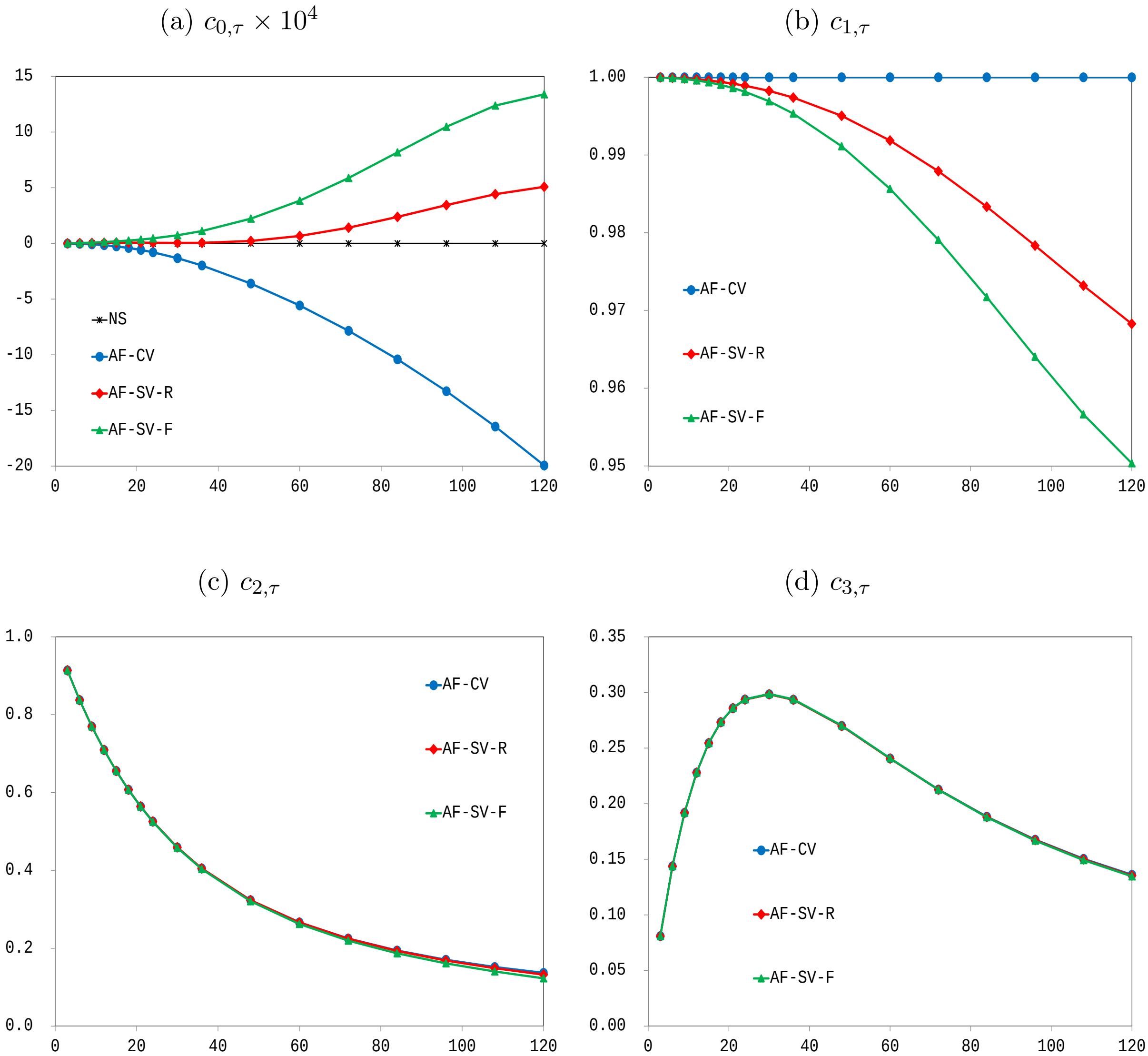

Figure 2: Coefficients of X_t in no-arbitrage yield function
The yield for the spanned-volatility model is linearly projected on a constant and X_t for each τ as $Y^{AFSV}(X_t, \tau) = c_{0,\tau} + c'_\tau X_t + u_{t,\tau}$, where $u_{t,\tau}$ contains remaining nonlinear terms of X_t. The OLS estimates of $c_{0,\tau}$ and c_τ are plotted against τ (month). For comparison, the corresponding plots for the NS model ($c_{0,\tau} = 0$ and $c_\tau = B(\tau)$) and the AF-CV model ($c_{0,\tau} = A(\tau)$ and $c_\tau = B(\tau)$) are presented. The parameter values used for computing the model-implied yields are given in Table 1 with $\lambda = 0.0609$.

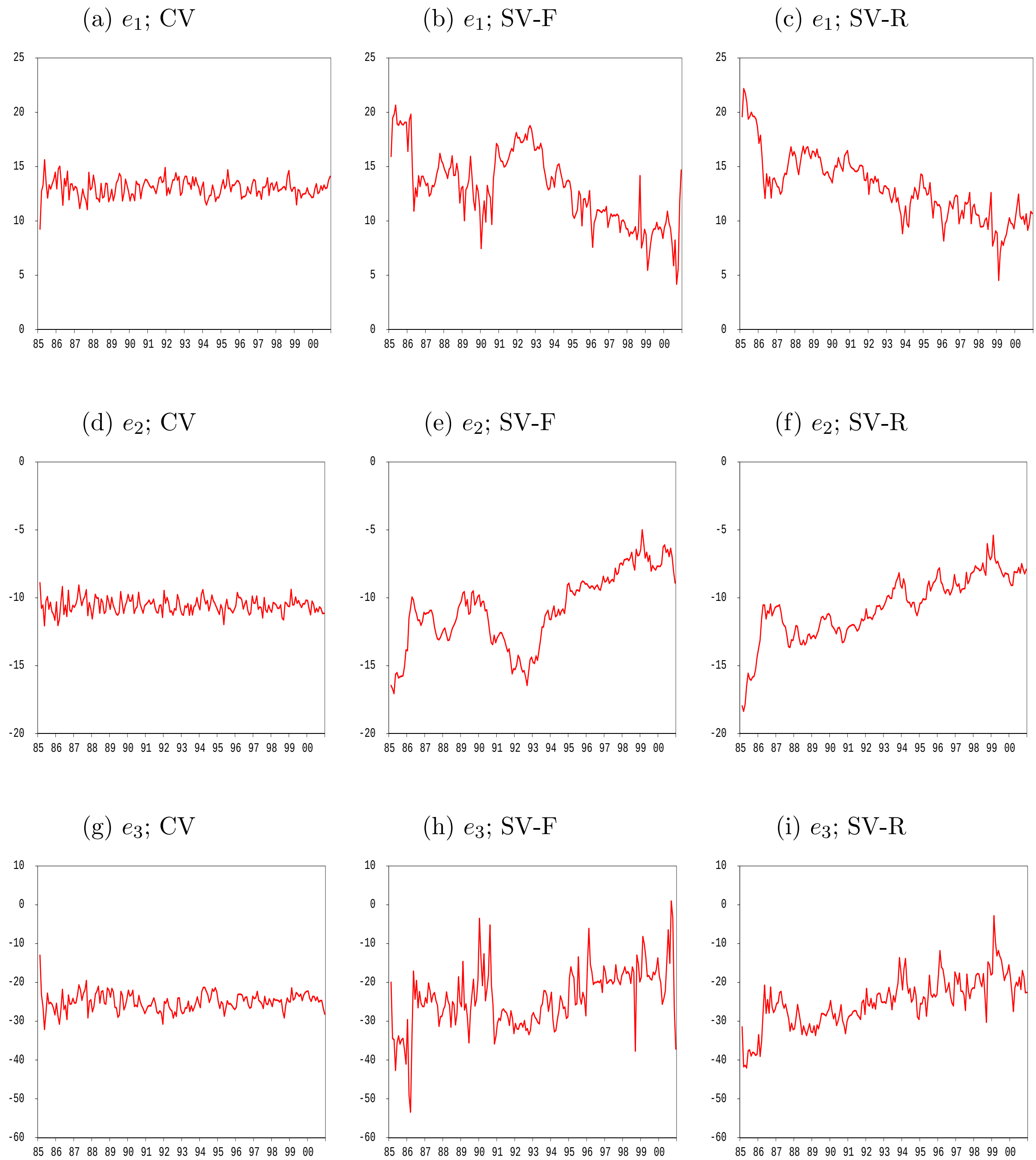

Figure 3: Time-series of difference in factors with and without no-arbitrage
The difference in each factor i is computed as $e^j_{i,t} = x^{AFj}_{i,t} - x^{NSj}_{i,t}$ $(i = 1, 2, 3;\ j = \{CV, SVF, SVR\})$. The factors are estimated by the one-step approach and the parameter values are given in Table 3 with $\lambda = 0.0609$. The sample period is from 1985/1 to 2000/12 (192 observations).

(a) AF-SV-F (b) AF-SV-R

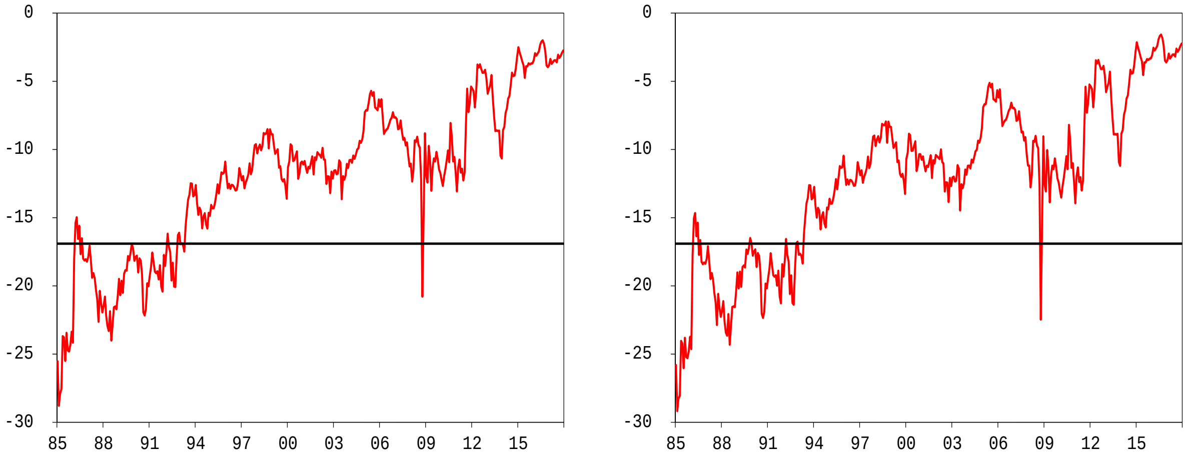

Figure B1: Time-series of difference in ten-year yields between spanned-volatility and NS models for GSW data
Figure B1 is analogous to Figure 1 of the main text, obtained with the dataset constructed by GSW (2007). The data are monthly, covering the period from 1985/1 to 2017/12 (396 observations).

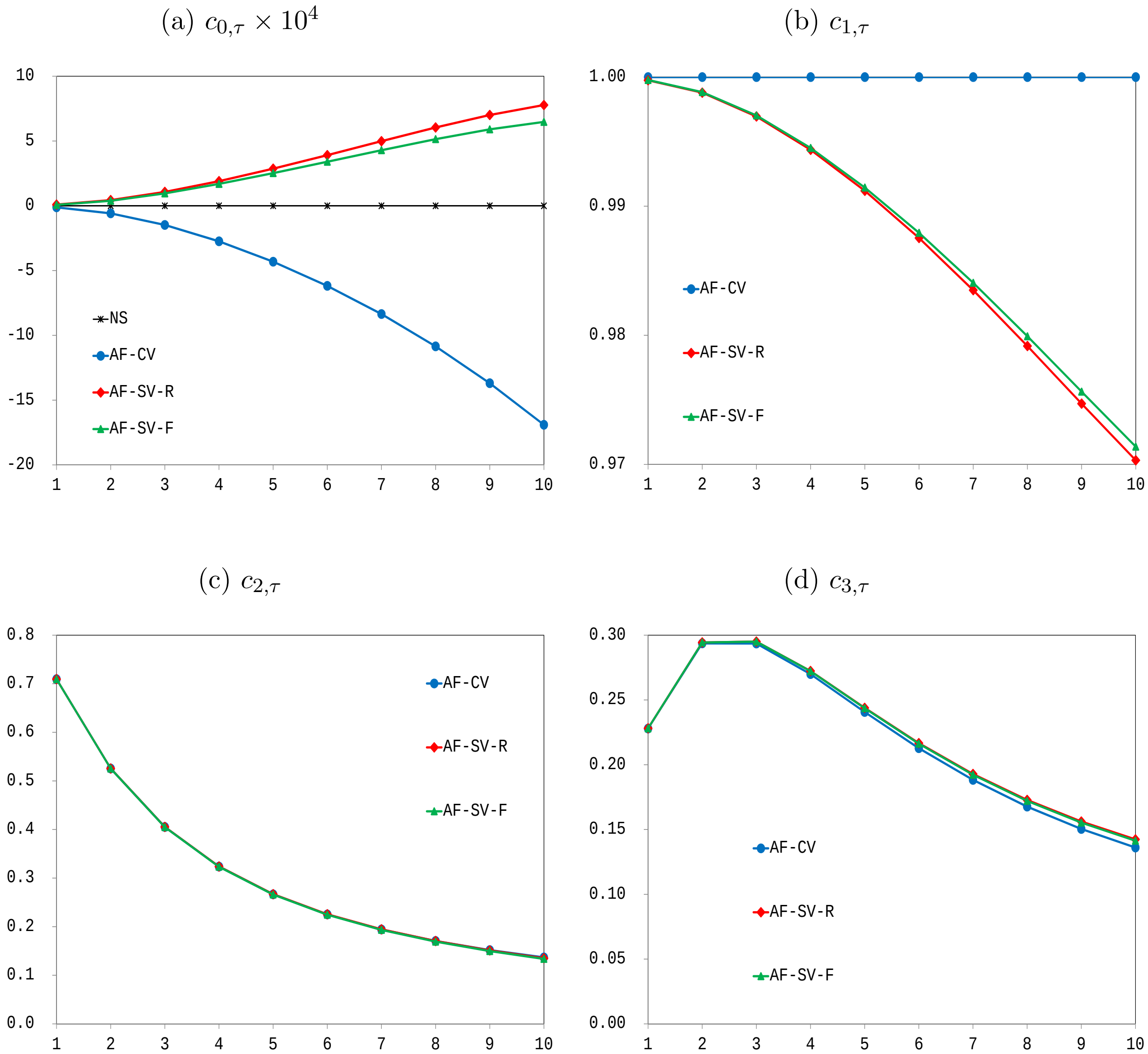

Figure B2: Coefficients of X_t in no-arbitrage yield function for GSW data
Figure B2 is analogous to Figure 2 of the main text, obtained with the dataset constructed by GSW (2007). The maturity ranges from one to ten years. The data are monthly, covering the period from 1985/1 to 2017/12 (396 observations).

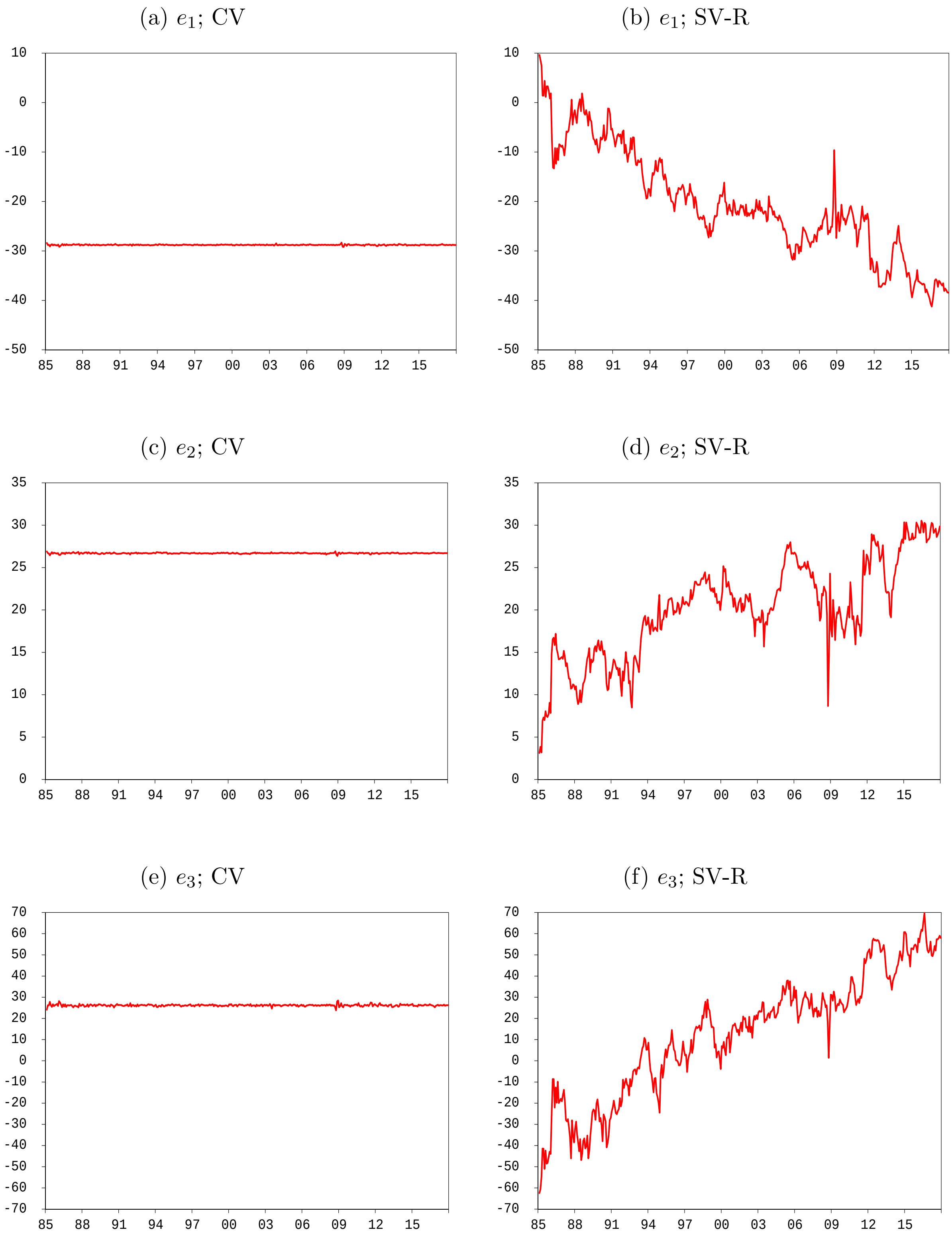

Figure B3: Time-series of difference in factors with and without no-arbitrage for GSW data

Figure B3 is analogous to Figure 3 of the main text, obtained with the dataset constructed by GSW. The data are monthly, covering the period from 1985/1 to 2017/12 (396 observations).

中央大学企業研究所
（The Institute of Business Research, Chuo University）
代表者　：　三浦　俊彦　（Director：Toshihiko MIURA）
〒192-0393　東京都八王子市東中野 742-1
（742-1 Higashinakano, Hachioji-shi, Tokyo 192-0393 JAPAN）
TEL　042-674-3272　（+81 42 674 3272）
FAX　042-674-3278　（+81 42 674 3278）
E-mail　kigyoken@tamajs.chuo-u.ac.jp
URL　http://www.chuo-u.ac.jp/research/institutes/business/