

Hevia, Constantino; Sola, Martin

Article

Bond risk premia and restrictions on risk prices

Journal of Risk and Financial Management

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Hevia, Constantino; Sola, Martin (2018) : Bond risk premia and restrictions on risk prices, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 11, Iss. 4, pp. 1-22,
<https://doi.org/10.3390/jrfm11040060>

This Version is available at:

<https://hdl.handle.net/10419/238928>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Article

Bond Risk Premia and Restrictions on Risk Prices [†]

Constantino Hevia and Martin Sola *

Department of Economics, Universidad Torcuato Di Tella, Buenos Aires 1428, Argentina; chevia@utdt.edu

* Correspondence: msola@utdt.edu

† We thank Agustín Gutierrez for outstanding research assistance.

Received: 18 August 2018; Accepted: 30 September 2018; Published: 3 October 2018



Abstract: Researchers who estimate affine term structure models often impose overidentifying restrictions (restrictions on parameters beyond those necessary for identification) for a variety of reasons. While some of those restrictions seem to have minor effects on the extracted factors and some measures of risk premia, such as the forward risk premium, they may have a large impact on other measures of risk premia that is often ignored. In this paper, we analyze how apparently innocuous overidentifying restrictions imposed on affine term structure models can lead to large differences in several measures of risk premiums.

Keywords: bond risk premia; affine term structure models; risk prices

JEL Classification: E43; G12

1. Introduction

Understanding bond risk premia remains one of the central issues in empirical finance. Much of this literature uses non-arbitrage affine term structure models (ATSM) to extract risk premium components from bond prices. The purpose of this paper is to analyze how apparently innocuous overidentifying restrictions imposed on ATSMs can lead to large differences in several measures of bond risk premia.

Long-term interest rates can be decomposed into expectations of average future short rates and a risk premium that investors demand for bearing long-term risk (the term premium). Forward interest rates can be split into an expected future interest rate and a risk premium component (the forward premium). Those are the measures of risk premia on which the literature has mostly focused. However, there are other definitions of risk premia that are arguably equally or more relevant from the point of view of market participants. An investor who buys a long-term bond financed by borrowing at a shorter interest rate is exposed to a risk premium (the bond holding risk premium). The return for an investor that writes a forward contract to buy a bond in the future, but closes the contract before the settlement date is also exposed to a risk premium (the holding forward risk premium). The contribution of this paper is to investigate to what extent different restrictions imposed on ATSMs, which often imply modest differences in the term and forward premiums, could lead to large differences in the bond holding and holding forward risk premiums.¹

Since often the extracted factors, forward premiums and term premiums look similar across restricted ATSM models, researchers tend to choose their preferred models based on the parsimony or forecasting ability of future yields. Yet, overidentifying restrictions unsupported by the data can have

¹ Although there is a large literature that has investigated the time variation in expected excess bond returns, it has mostly been done in the context of documenting the failure of the expectations hypothesis. This literature is too large to summarize here, but some widely-cited papers are Fama and Bliss (1987), Campbell and Shiller (1991) and Cochrane and Piazzesi (2005).

a large impact on other measures of risk premia, such as the bond holding risk premium or the holding forward risk premium, particularly at longer horizons. The reason is that small errors compound and get amplified as the holding horizon increases. Therefore, if the objective of the analysis is to understand bond risk premia, researchers should be very careful in not imposing overidentifying restrictions unsupported by the data even though the more parsimonious model may produce similar factors or predict better than the unrestricted model. Those restrictions can be harmful.

The main, and perhaps only, reason to consider affine models of bond prices is to understand risk premia. As noted by [Joslin et al. \(2011\)](#) and [Duffee \(2011\)](#), if the objective of the analysis is to fit a yield curve or to forecast bond yields, little is gained by imposing non-arbitrage restrictions. However, to understand risk premia, non-arbitrage restrictions are of fundamental importance. Unfortunately, there is no clear guidance in the literature on how to impose restrictions on risk prices. This issue has been overlooked in the literature, and this paper is an attempt to fill this gap. To our knowledge, [Bauer \(2018\)](#) is the only paper that studied restrictions on risk prices. He did so from a Bayesian perspective and only focusing on the term premium and excess bond returns. This paper is complementary to Bauer's in that we also analyze restrictions on risk prices, but from a different perspective. In particular, besides the term premium and excess returns, we analyze how restrictions on risk prices affect other measures of risk premia.

We first analyze to what extent usual restrictions imposed on risk prices or factor dynamics imply similar estimated term and forward premiums, but may lead to large differences in excess holding returns and holding forward premiums. The different concepts of risk premia are related among them, and small deviations in estimated forward premiums accumulate and can lead to large differences in expected excess returns of long-term bonds. For example, the bond holding risk premium is proportional to the forward risk premium, where the factor of proportionality increases with the maturity of the bond. Small differences in the estimated one-month forward risk premium, which may seem irrelevant if one focuses only on the forward premium, get amplified by about 120 times when we compute the risk premium associated with holding a 10-year bond financed by borrowing at the one-month interest rate.

Although the argument is general, we base our empirical investigation using the class of arbitrage-free Nelson and Siegel models. The work in [Nelson and Siegel \(1987\)](#) proposed a flexible, yet parsimonious functional form to model the cross-section of bond yields, which is widely used by practitioners and researchers.² The work in [Christensen et al. \(2011\)](#) showed that the Nelson and Siegel parametric representation of the yield curve can be made arbitrage-free by including a maturity-specific constant to the traditional Nelson and Siegel model. We use the Nelson and Siegel model because it is easy to estimate, and it is sufficiently general to make our point. We analyze several restrictions on risk prices and factor dynamics that are considered in the literature to analyze risk premia. Although the estimated factors are virtually identical across models, the estimates of risk premia that they produce often vary dramatically because the loadings on those factors change across models.

Using as a baseline the most general arbitrage-free Nelson and Siegel model, we evaluate to what extent the econometric rejections of the restrictions on risk prices are economically relevant (in terms of the estimated risk premia). For example, imposing a diagonal covariance matrix on the evolution of the risk factors is comfortably rejected using standard econometric tests, but the estimated risk premia are similar to those of the unrestricted model. On the other hand, imposing a diagonal matrix on the lagged values of the vector autoregression representation of the risk factors is also rejected, but with a much smaller likelihood ratio statistic. Yet, the estimated risk premiums in the restricted model are very different from those of the unrestricted model.

² For example, the arbitrage-free Nelson and Siegel model has been used in [Christensen et al. \(2010\)](#), [Gürkaynak and Wright \(2012\)](#), [Christensen and Rudebusch \(2015\)](#), among many others. The reference [Diebold and Rudebusch \(2012\)](#) is a textbook treatment of the dynamic Nelson and Siegel model.

The main practical implication of these results is that, to analyze and forecast risk premia, it is advisable to consider the just identified model. Imposing restrictions that are unsupported by the data may lead to large errors in estimated and forecast measures of risk premia. These observations also have policy implications: to the extent that fluctuations in risk premia have macroeconomic consequences, a policy maker that reacts to spurious changes in risk premia may choose suboptimal policies.

Using an ex-ante empirical analysis, [Cochrane and Piazzesi \(2005, 2008\)](#) argued that excess returns are a function of a single factor and built an affine term structure model imposing this assumption on risk prices. Other papers used likelihood ratios to test whether certain risk prices are zero, e.g., ([Joslin et al. 2011](#)). Some restrictions are often motivated on forecasting grounds. For example, [Christensen et al. \(2011\)](#) argued that imposing a diagonal structure on the factor dynamics outperforms a more general model in forecasting the yield curve. Even though this may be the correct procedure when the objective of the exercise is to forecast yields, it may be counterproductive when the objective is to understand risk premia. The reason is well known: even if the “true model” may contain many parameters, it may be outperformed out-of-sample by invalid reductions in the number of free parameters due to parameter uncertainty. When the issue is to understand the evolution and determinants of risk premia, restrictions that are not supported by the data using standard testing procedures may have damaging effects on the evaluation of risk compensation.³

The paper is organized as follows. Section 2 describes a general affine term structure model and the four different concepts of risk premia that we consider. Section 3 provides a Nelson–Siegel representation of the affine model that is used in the empirical section of the paper. Section 4 shows the estimation results, and Section 5 analyzes the impact of imposing restrictions on risk prices. Section 6 concludes, and the Appendix contains some proofs.

2. An Affine Model of Bond Prices

Let X_t denote a $k \times 1$ vector of unobserved risk factors that summarizes the information that investors use to price discount bonds. The risk factors evolve as a first order vector autoregression:

$$X_{t+1} = \mu + \Phi X_t + \Gamma u_{t+1}, \tag{1}$$

where Γ is lower triangular and $u_{t+1} \sim \text{i.i.d.} N(0, I_k)$.

Investors price nominal cash flows using the one-period stochastic discount factor:

$$\ln M_{t+1} = -\delta_0 - \delta_1' X_t - \frac{1}{2} \Lambda_t' \Lambda_t - \Lambda_t' u_{t+1}. \tag{2}$$

The $k \times 1$ vector Λ_t , referred to as the market price of risk, represents the compensation that investors demand to face shocks to the state vector u_{t+1} . The market price of risk is an affine function of the factors:

$$\Lambda_t = \lambda_0 + \lambda_1 X_t, \tag{3}$$

where λ_0 is a $k \times 1$ vector and λ_1 is a $k \times k$ matrix of coefficients.

The principle of no-arbitrage implies that bond prices satisfy the recursion:

$$P_t^{(n)} = E_t(M_{t+1} P_{t+1}^{(n-1)}). \tag{4}$$

³ One possible concern with the methodology used in this paper is that the 2008–2009 global crisis was characterized by significant asymmetries in the relation between yield spreads and macroeconomic factors ([Evgenidis et al. 2017](#)). Clearly, the affine model is unsuitable to study asymmetric responses to the business cycle. Yet, as shown in [Hevia et al. \(2015\)](#), it is possible to extend the affine model of bond yields to allow for discrete changes in the state of the economy using a Markov switching framework, thus capturing asymmetric responses during booms and recessions. The study of restrictions on risk prices in such a context; however, it is beyond the scope of this paper.

As shown by [Ang and Piazzesi \(2003\)](#), log bond prices, $p_t^{(n)} = \ln P_t^{(n)}$, are also an affine function of the risk factors:

$$p_t^{(n)} = A_n + B_n' X_t, \tag{5}$$

where the coefficients A_n and B_n satisfy the recursions:

$$B_{n+1} = \Phi^{*'} B_n - \delta_1, \tag{6}$$

$$A_{n+1} = A_n + \mu^{*'} B_n + \frac{1}{2} B_n' \Gamma \Gamma' B_n - \delta_0, \tag{7}$$

with $A_0 = 0, B_0 = 0$ and where μ^* and Φ^* are defined as:

$$\begin{aligned} \mu^* &= \mu - \Gamma \lambda_0 \\ \Phi^* &= \Phi - \Gamma \lambda_1. \end{aligned}$$

The continuously-compounded yield on an n -period discount bond is thus:

$$y_t^{(n)} = -p_t^{(n)}/n = a_n + b_n' X_t, \tag{8}$$

with $a_n = -A_n/n$ and $b_n = -B_n/n$.

We are also interested in the expected returns from holding forward contracts on discount bonds. Let $F_t^{(s,n-s)}$ denote the settlement price of a forward contract entered into at time t to buy an $(n-s)$ -period discount bond at time $t+s$, with $n > s$. Since two different ways of moving one dollar from t to $t+n$ must cost the same, the law of one price implies:

$$P_t^{(n)} = P_t^{(s)} F_t^{(s,n-s)}. \tag{9}$$

The log forward rate $f_t^{(s,n-s)}$ is the implicit continuously-compounded interest rate of a forward contract with settlement price $F_t^{(s,n-s)}$,

$$F_t^{(s,n-s)} = e^{-(n-s)f_t^{(s,n-s)}}.$$

In terms of bond prices, the log forward rate $f_t^{(s,n-s)}$ is thus:

$$f_t^{(s,n-s)} = \frac{1}{n-s} [p_t^{(s)} - p_t^{(n)}]. \tag{10}$$

Therefore, the bond prices (5) imply that forward rates are affine functions of the factors,

$$f_t^{(s,n-s)} = a_{s,n}^f + b_{s,n}^{f'} X_t,$$

where $a_{s,n}^f = (A_s - A_n)/(n-s)$ and $b_{s,n}^{f'} = (B_s - B_n)/(n-s)$.

In what follows, we use different concepts of returns, all of which are measured on a monthly basis. The h -period log holding return from buying an n -period discount bond at time t and selling it as an $(n-h)$ -period bond at time $t+h$ is:

$$r_{b,t \rightarrow t+h}^{(n)} = \frac{1}{h} (p_{t+h}^{(n-h)} - p_t^{(n)}).$$

Denote the h -period log return from holding the n -period bond in excess of the yield of an h -period bond by:

$$rx_{b,t \rightarrow t+h}^{(n)} = r_{b,t \rightarrow t+h}^{(n)} - y_t^{(h)}.$$

Suppose now that, at time t , an investor enters into a forward contract to buy at time $t + s$ an $(n - s)$ -period discount bond, but closes the position at time $t + h < t + s$. Closing the position is equivalent to writing a forward contract at time $t + h$ on an $(n - s)$ -period bond with settlement date $t + s$. We define the h -period log holding return from holding the proposed forward as the difference between the log-forward rates associated with each of the contracts,

$$r_{f,t \rightarrow t+h}^{(s,n-s,h)} = f_t^{(s,n-s)} - f_{t+h}^{(s-h,n-s)}.$$

2.1. Bond Risk Premia

Our goal is to analyze how different restrictions on risk prices affect bond risk premia. However, what risk premium? Risk premia in bond markets are usually defined in terms of deviations from the expectational hypothesis. Four equivalent forms of the expectational hypothesis give rise to different, but related concepts of risk premia,⁴

1. The forward premium: The time t forward rate for loans between times $t + s$ and $t + n$ is the expected $(n - s)$ -period yield at time $t + s$ plus a risk premium,

$$f_t^{(s,n-s)} = E_t \left[y_{t+s}^{(n-s)} \right] + \pi_{f,t}^{(s,n-s)}. \tag{11}$$

2. The bond holding risk premium: The expected h -period log holding return of an n -period bond measured on a monthly rate is the yield of an h -period bond plus a risk premium:

$$E_t [r_{b,t \rightarrow t+h}^{(n)}] = y_t^{(h)} + \pi_{hb,t}^{(n)}. \tag{12}$$

3. The term premium: The n -period yield is the average of expected one-period yields plus a risk premium,

$$y_t^{(n)} = \frac{1}{n} \left(\sum_{j=0}^{n-1} E_t y_{t+j}^{(1)} \right) + \pi_{y,t}^{(n)}. \tag{13}$$

4. The holding forward risk premium: Since buying or writing a forward contract costs zero, any expected return from holding a forward contract is a risk premium,

$$E_t \left[r_{f,t \rightarrow t+h}^{(s,n-s,h)} \right] = \pi_{hf,t}^{(s,n-s,h)}, \tag{14}$$

for all $h, s > h$, and $n > s + h$.

The four definitions of risk premia are equivalent in the sense that if one of them is zero or constant, so are the other three (see the Appendix A). For our purposes, it is useful to express the risk premiums $\pi_{y,t}^{(n)}$, $\pi_{hb,t}^{(n,h)}$ and $\pi_{hf,t}^{(s,n-s,h)}$ as a function of the forward premium $\pi_{f,t}^{(j,1)}$,

$$\begin{aligned} \pi_{y,t}^{(n)} &= \frac{1}{n} \sum_{j=1}^h \pi_{f,t}^{(j,1)}, \\ \pi_{hf,t}^{(s,n-s,h)} &= \pi_{f,t}^{(s,n-s)} - E_t \left[\pi_{f,t+h}^{(s-h,n-s)} \right], \\ \pi_{hb,t}^{(n,h)} &= \frac{n-h}{h} \pi_{f,t}^{(h,n-h)}. \end{aligned}$$

⁴ The work in [Cochrane and Piazzesi \(2008\)](#) considered the definitions of Risk Premia 1 through 3 below. We add the fourth, which states that all expected log holding forward returns are risk premia.

In particular, the bond holding risk premium $\pi_{hb,t}^{(n,h)}$ is proportional to the forward premium $\pi_{f,t}^{(h,n-h)}$, where the factor of proportionality increases with the maturity n of the bond for a given holding period h . Small differences in the estimated forward premium due to invalid parameter restrictions, which may seem irrelevant to the naked eye, accumulate and may drive large and economically meaningful differences in the bond holding risk premium. For example, a 10 basis point difference in the one-month forward premium ($h = 1$) can lead to differences as large as 12 percentage points in the estimated one-month bond holding risk premium associated with holding a 10-year bond ($n = 120$) financed by borrowing at the one-month rate. In this case, the factor of proportionality is $(n - h)/h = 119$.

The closed form expressions of the risk premiums using the structure of the affine model are the following,

$$\pi_{f,t}^{(h,n-h)} = J_f^{(h,n-h)} + \frac{B'_{n-s}}{n-s} \left[\mu^{(s-1)} - \mu^{*(s-1)} + (\Phi^s - \Phi^{*s}) X_t \right], \tag{15}$$

$$\pi_{hb,t}^{(n,h)} = J_{hb}^{(n,h)} + \frac{B'_{n-h}}{h} \left[\mu^{(h-1)} - \mu^{*(h-1)} + (\Phi^h - \Phi^{*h}) X_t \right], \tag{16}$$

$$\pi_{y,t}^{(n)} = J_y^{(n)} + \frac{B'_1}{n} \sum_{j=0}^{n-1} \left[\mu^{(j-1)} - \mu^{*(j-1)} + (\Phi^j - \Phi^{*j}) X_t \right], \tag{17}$$

$$\pi_{hf,t}^{(s,n-s,h)} = J_{hf}^{(s,n-s,h)} + \frac{B'_{n-h} - B'_{s-h}}{n-s} \left[\mu^{(h-1)} - \mu^{*(h-1)} + (\Phi^h - \Phi^{*h}) X_t \right], \tag{18}$$

where we define $\mu^{(j)} = (I + \Phi + \Phi^2 + \dots + \Phi^j) \mu$; $\mu^{*(j)} = (I + \Phi^* + \Phi^{*2} + \dots + \Phi^{*j}) \mu^*$; and $J_f^{(h,n-h)}$, $J_{hb}^{(n,h)}$, $J_y^{(n)}$ and $J_{hf}^{(s,n-s,h)}$ are time invariant Jensen inequality terms.⁵ Note that with risk neutrality (when the physical and risk neutral measure coincide, $\Phi = \Phi^*$), risk premia collapse to the usual Jensen inequality terms.

Even though different restrictions on the parameters of the affine model usually deliver similar estimated risk factors X_t , the evolution of risk premia may vary across these models as long as the matrices Φ and Φ^* are different. Furthermore, for a given holding period h , the bond holding risk premium $\pi_{hb,t}^{(n,h)}$ is more sensitive to differences between Φ and Φ^* than the forward risk premium the longer is the maturity of the bond n . For example, when $h = 1$, the loading of the holding bond risk premia on X_t is $B'_{n-1}(\Phi - \Phi^*)$, while that of the forward premium is $B'_{n-1}(\Phi - \Phi^*)/(n - 1)$. Thus, for the bond with the longest maturity in our sample, $n = 120$, the loading of the forward premium on X_t is almost 120 times smaller than that of the bond holding risk premium. Therefore, small differences in the estimated coefficients may lead to very different estimates of bond holding risk premiums for long maturity bonds even though the forward premiums may look similar.⁶

⁵ The Jensen inequality terms are given by:

$$\begin{aligned} J_f^{(h,n-h)} &= \sum_{j=0}^{s-1} \frac{1}{n-s} \left(B'_j \frac{\Gamma'}{2} B_j - B'_{n-s+j} \frac{\Gamma'}{2} B_{n-s+j} \right) \\ J_{hb}^{(n,h)} &= \sum_{j=0}^{h-1} \frac{1}{h} \left(B'_j \frac{\Gamma'}{2} B_j - B'_{n-h+j} \frac{\Gamma'}{2} B_{n-h+j} \right) \\ J_y^{(n)} &= -\frac{1}{n} \sum_{j=1}^{n-1} B'_j \frac{\Gamma'}{2} B_j \\ J_{hf}^{(s,n-s,h)} &= \sum_{i=0}^{n-s-1} \frac{1}{n-s} \left(B'_{s-h+i} \frac{\Gamma'}{2} B_{s-h+i} - B'_{s+i} \frac{\Gamma'}{2} B_{s+i} \right). \end{aligned}$$

Proving Expressions (15) through (18) is not particularly difficult, but it is lengthy and tedious. We provide detailed proofs upon request.

⁶ The result is reversed if $n/h < h$ (so that $n - h < h$).

3. A Nelson–Siegel Representation of the Affine Model

General unrestricted affine models with unobserved factors are unidentified because the factors can be re-scaled, rotated and translated without affecting the empirical implications of the model. There are several representations of affine models that are identified. We follow [Christensen et al. \(2011\)](#) and use a representation of the arbitrage-free model in which the yield curve adopts an augmented Nelson–Siegel parametrization of the cross-section of bond yields. As noted by [Christensen et al. \(2011\)](#) and [Hamilton and Wu \(2012\)](#), the Nelson–Siegel model is a convenient representation that is econometrically identified (subject to minor identification restrictions) and that is easy to estimate using the method of maximum likelihood. This representation of the ATSM is sufficiently general to make our point that imposing invalid overidentifying restrictions on risk prices, which may look irrelevant from the point of view of some measures of risk premiums, could lead to large differences in other measures of risk premia.

The dynamic Nelson–Siegel model is a three-factor model that fits the cross-section and time series of yield remarkably well ([Diebold and Li 2006](#)). Yet, in its traditional representation, the Nelson–Siegel model does not rule out arbitrage opportunities. The work in [Christensen et al. \(2011\)](#) showed that one can augment the Nelson–Siegel model with a maturity-specific constant to render it arbitrage-free. In particular, the yield $y_t^{(n)}$ at any time t as a function of the maturity of the bond n is parametrized as:

$$y_t^{(n)} = a_n + \zeta_{1t} + \left(\frac{1 - e^{-\theta n}}{\theta n}\right) \zeta_{2t} + \left(\frac{1 - e^{-\theta n}}{\theta n} - e^{-\theta n}\right) \zeta_{3t}, \tag{19}$$

where $(\zeta_{1t}, \zeta_{2t}, \zeta_{3t})'$ are latent variables interpreted as level, slope and curvature factors; and the parameter θ determines the shape of the loadings on the factors.

In the traditional Nelson–Siegel model, $a_n = 0$ for all n . In the arbitrage-free Nelson–Siegel model, the vector of risk factors is $X_t = [\zeta_{1t}, \zeta_{2t}, \zeta_{3t}]'$, and the factor loadings and constant a_n are chosen to satisfy the recursions (6) and (7). As shown by [Christensen et al. \(2011\)](#) in a continuous time setting, rendering the Nelson–Siegel model arbitrage-free requires imposing restrictions on the risk-neutral evolution of the state variables, $X_{t+1} = \mu^* + \Phi^* X_t + \Gamma \eta_{t+1}^*$, where $\mu^* = \mu - \Gamma \lambda_0$, $\Phi^* = \Phi - \Gamma \lambda_1$ and $\eta_{t+1}^* \sim N(0, I)$. The conditions are as follows: consider any 3×1 vector μ^* and a risk-neutral matrix on lagged values Φ^* defined as:

$$\Phi^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\theta} & \theta e^{-\theta} \\ 0 & 0 & e^{-\theta} \end{bmatrix}, \tag{20}$$

where $\theta > 0$. Furthermore, let δ_0 be any number and $\delta_1' = \left[1, \frac{1 - e^{-\theta}}{\theta}, \frac{1 - e^{-\theta}}{\theta} - e^{-\theta}\right]$. Using these assumptions, the recursions (6) imply:

$$b_n = -B_n/n = \left[1, \frac{1 - e^{-\theta n}}{\theta}, \frac{1 - e^{-\theta n}}{\theta n} - e^{-\theta n}\right]', \tag{21}$$

precisely the factor loadings in the Nelson–Siegel representation (19). The maturity-specific constant a_n is set to satisfy (7). Given these values, the parameters of the market price of risk λ_0 and λ_1 are given by $\lambda_0 = \Gamma^{-1}(\mu - \mu^*)$ and $\lambda_1 = \Gamma^{-1}(\Phi - \Phi^*)$.

Identification of the Nelson–Siegel model requires imposing simple restrictions on the parameter δ_0 and the risk-neutral intercept μ^* . As in [Dai and Singleton \(2000\)](#), we set $\delta_0 = 0$. Furthermore, since in the recursion (7), $\mu^{*'} B_n$ is a scalar, we can identify a single parameter in μ^* . We set $\mu^{*'} = [\mu_1^*, 0, 0]$ and estimate μ_1^* as a free parameter along with the other parameters of the model.

Let $\mathcal{N} > 3$ denote the set of yield maturities observed by the econometrician. Since there are more observed maturities than factors, the model is stochastically singular. One possibility is to assume that three yields, or three linear combinations of yields, are perfectly priced by the model and then to impose

classical measurement errors into the remaining yields (see (Hamilton and Wu 2012; Joslin et al. 2011) among others). In contrast, the Nelson–Siegel literature, e.g., (Christensen et al. 2011), assumes that all yields are measured with error, which is an appealing assumption to make since which yields are perfectly priced or measured without error is arbitrary. We follow the latter approach and assume that all yields are observed with uncorrelated measurement errors. Thus, for each yield $y_t^{(n)}$ with $n \in \mathcal{N}$, the arbitrage-free Nelson–Siegel model is represented by the following system of equations:

$$y_t^{(n)} = a_n + \zeta_{1t} + \frac{1-e^{-\theta n}}{\theta} \zeta_{2t} + \left(\frac{1-e^{-\theta n}}{\theta n} - e^{-\theta n} \right) \zeta_{3t} + v_t^{(n)}, \tag{22}$$

$$X_{t+1} = \mu + \Phi X_t + \Gamma \eta_{t+1}, \tag{23}$$

where the state vector $X_t = [\zeta_{1t}, \zeta_{2t}, \zeta_{3t}]'$ is unobserved, the intercept a_n satisfies the recursion (7) and $v_t^{(n)} \sim \text{i.i.d. } N(0, \sigma_n^2)$.

Since the risk factors are unobserved, and it is not possible to recover X_t from the observed yields, we use the Kalman filter to evaluate the prediction-error decomposition of the likelihood function. In the unrestricted Nelson–Siegel model, we maximize the log-likelihood function numerically choosing the free parameters $\mu, \Phi, \Gamma, \mu_1^*, \theta$ and σ_n^2 for $n \in \mathcal{N}$. In restricted versions of the Nelson–Siegel model, we impose the appropriate restrictions on the free parameters.

4. Estimation Results

We used data on U.S. Treasury yields of fixed maturities of 3, 6, 12, 24, 36, 48, 60, 72, 84, 96, 108 and 120 months from January 1985 through December 2013. The shortest yield is the three-month treasury constant maturity rate obtained from the Federal Reserve Bank of St. Louis (Series DGS3MO). The remaining yields are from Cochrane (2015), who updated the data in Joslin et al. (2011) until December 2013.

We begin by estimating the unrestricted unobserved factor model imposing the arbitrage-free Nelson–Siegel parametrization discussed above. Using this model as a benchmark, we imposed restrictions commonly used in the literature on the parameters that determined the evolution of the risk factors under the physical measure. Those restrictions were often motivated in terms of forecasting performance or simplicity grounds. In particular, in the baseline case, we did not restrict any of the parameters of the model besides the aforementioned identifying restrictions. Since all the restricted models that we consider are nested, we next tested whether the restrictions were supported by the data using likelihood-ratio tests along with information criteria for in-sample comparisons. More importantly, we assessed the economic relevance of the restrictions by investigating their impact on estimated risk premia.

Table 1 shows the estimation results of the general model along with the log-likelihood value, information criteria and estimated parameter θ of the different restricted models. The general model corresponds to the Nelson–Siegel parametrization without imposing any restriction on the evolution of the state variables. The restricted models impose different constraints on the state equation: we considered a model with a diagonal Φ matrix, as in Christensen et al. (2011); a model with a diagonal Γ matrix, as in Gürkaynak and Wright (2012); and a model in which only shocks that affect the market price of level risk $\Lambda_{1,t}$ (the first element of Λ_t) are priced, in the spirit of Cochrane and Piazzesi (2005, 2008). Under this last parametrization, the matrix λ_1 had non-zeros only in the first row, and only the first element of the vector λ_0 was non-zero.⁷

⁷ Since Cochrane and Piazzesi (2005, 2008) developed a four-factor model (level, slope, curvature factors plus a return forecasting) factor, their results and ours are not exactly comparable. Yet, after some pre-estimation analysis, they argued that only shocks to their forecasting factor affected the market price of level risk $\Lambda_{1,t}$. Since the forecasting factor captures information of the level, slope and curvature of the yield curve (and possibly some other information), we captured their restrictions on risk prices by allowing shocks to the three factors to affect only the market price of level risk and set all the other elements of the matrix λ_1 (the second and third rows) to zero.

Table 1. Estimates of arbitrage-free Nelson and Siegel models.

Unrestricted Arbitrage-Free Nelson and Siegel Model						
<i>Parameters of the VAR(1) process for the yield curve factors ($\zeta_{1t}, \zeta_{2t}, \zeta_{3t}$)</i>						
$\mu (\times 1000)$	Φ			$\Gamma (\times 1000)$		
$\begin{bmatrix} 0.1654 \\ (0.0630) \\ 0.0079 \\ (0.0716) \\ -0.4064 \\ (0.1633) \end{bmatrix}$	$\begin{bmatrix} 0.967 \\ (0.014) \\ -0.016 \\ (0.016) \\ 0.096 \\ (0.033) \end{bmatrix}$	$\begin{bmatrix} -0.006 \\ (0.015) \\ 0.960 \\ (0.016) \\ 0.055 \\ (0.034) \end{bmatrix}$	$\begin{bmatrix} 0.023 \\ (0.011) \\ 0.027 \\ (0.013) \\ 0.901 \\ (0.028) \end{bmatrix}$	$\begin{bmatrix} 0.274 \\ (0.013) \\ -0.247 \\ (0.016) \\ -0.120 \\ (0.045) \end{bmatrix}$	$\begin{bmatrix} 0 \\ (0.008) \\ 0.200 \\ (0.008) \\ 0.035 \\ (0.038) \end{bmatrix}$	$\begin{bmatrix} 0 \\ (0.008) \\ 0 \\ (0.008) \\ 0.678 \\ (0.029) \end{bmatrix}$
<i>Other parameters</i>						
$\theta = 0.0403 (0.001) \quad \mu_1^Q (\times 1000) = 0.0094 (0.0009)$						
Log-likelihood = 33,225.3 Akaike = -66,386.5 Schwarz = -66,263.2						
Model with diagonal Φ matrix						
Log-likelihood = 33,200.7 Akaike = -66,349.4 Schwarz = -66,249.2 $\theta = 0.0402$						
LR statistic = 49.14 p-value = 0.000						
Model with diagonal Γ matrix						
Log-likelihood = 3,084.4 Akaike = -66,110.9 Schwarz = -65,999.2 $\theta = 0.0407$						
LR statistic = 281.61 p-value = 0.000						
Model with only market price of level risk						
Log-likelihood = 33,185.0 Akaike = -66,322.0 Schwarz = -66,229.6 $\theta = 0.0401$						
LR statistic = 80.47 p-value = 0.000						

The likelihood ratio tests and the information criteria suggested that none of the models were valid restrictions of the general model. Yet, the estimated risk factors X_t and the parameter θ , which governed the evolution of the risk factors under the risk-neutral measure, were virtually the same in all models (Figure 1 and Table 1). Since bond prices depend only on the parameters of the risk neutral measure (determined only by θ and μ_1^* in the Nelson–Siegel parametrization), the four models provided a virtually identical characterization of the cross-section of bond yields. Risk premiums, however, depend on the evolution of the risk factors under the physical measure, which did vary in the different models. Alternatively, the estimated values of the parameters λ_0 and λ_1 , displayed in Table 2, differed substantially across models, often even with different signs.

Table 2. Estimated market prices of risk parameters.

Unrestricted model				Diagonal Φ model			
$\Gamma\lambda_0 (\times 1000)$	$\Gamma\lambda_1$			$\Gamma\lambda_0 (\times 1000)$	$\Gamma\lambda_1$		
$\begin{bmatrix} 0.156 \\ 0.008 \\ -0.406 \end{bmatrix}$	$\begin{bmatrix} -0.0333 \\ -0.0159 \\ 0.0958 \end{bmatrix}$	$\begin{bmatrix} -0.0063 \\ -0.0004 \\ 0.0553 \end{bmatrix}$	$\begin{bmatrix} 0.0233 \\ -0.0119 \\ -0.0592 \end{bmatrix}$	$\begin{bmatrix} 0.010 \\ -0.034 \\ -0.033 \end{bmatrix}$	$\begin{bmatrix} -0.0030 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0.0285 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ -0.0386 \\ -0.0064 \end{bmatrix}$
Diagonal Γ model				Only market price of level risk			
$\Gamma\lambda_0 (\times 1000)$	$\Gamma\lambda_1$			$\Gamma\lambda_0 (\times 1000)$	$\Gamma\lambda_1$		
$\begin{bmatrix} 0.091 \\ 0.068 \\ -0.359 \end{bmatrix}$	$\begin{bmatrix} -0.0184 \\ -0.0266 \\ 0.0841 \end{bmatrix}$	$\begin{bmatrix} 0.0009 \\ -0.0066 \\ 0.0628 \end{bmatrix}$	$\begin{bmatrix} 0.0116 \\ -0.0001 \\ -0.0617 \end{bmatrix}$	$\begin{bmatrix} 0.135 \\ -0.115 \\ -0.062 \end{bmatrix}$	$\begin{bmatrix} -0.0273 \\ 0.0233 \\ 0.0125 \end{bmatrix}$	$\begin{bmatrix} 0.0009 \\ -0.0008 \\ -0.0004 \end{bmatrix}$	$\begin{bmatrix} 0.0150 \\ -0.0128 \\ -0.0068 \end{bmatrix}$

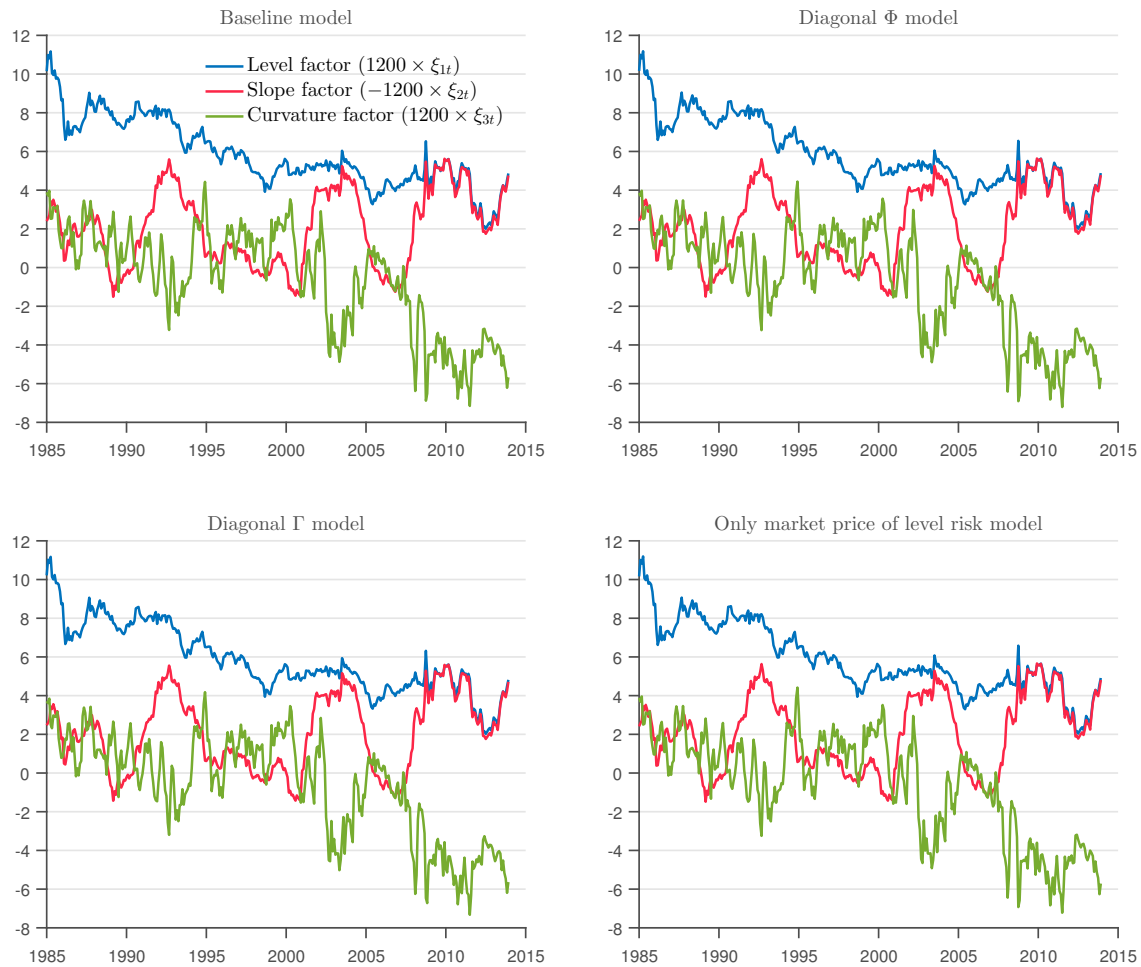


Figure 1. Estimated unobserved factors and empirical level, slope and curvature.

5. Risk Premia and Restrictions on Risk Prices

In this section, we analyze how imposing different restrictions on the general model affects the four measures of risk premium discussed above.

5.1. The Forward Risk Premium

The forward risk premium can be computed from forecasts of the future yield curve (Equation (11)). To the extent that these forecasts are similar to those obtained from reduced form vector autoregression, the forward risk premium would be almost model independent (see Joslin et al. 2011, Hamilton and Wu 2012 and Duffee (2011)).

In Figure 2, we show the estimated forward premiums in the four versions of the model. The upper left panel displays the premium associated with opening a forward contract with a settlement date one year ahead ($s = 1$ using our notation) on a bond that matures in one year ($n - s = 1$). In general, the notation s -for- $(n - s)$ forward represents the premium associated with opening a forward contract with settlement date s years ahead on a bond that matures in $n - s$ years. Except for the factor model in which only the market price of level risk matters, the forward premium is roughly similar across models, particularly for longer term bonds. Although not shown in the figure, all forward premiums move almost one-to-one with the slope factor ζ_{2t} of the yield curve. In the following subsections, we show how these small differences in the forward premium across models get amplified when considering other definitions of risk premium.

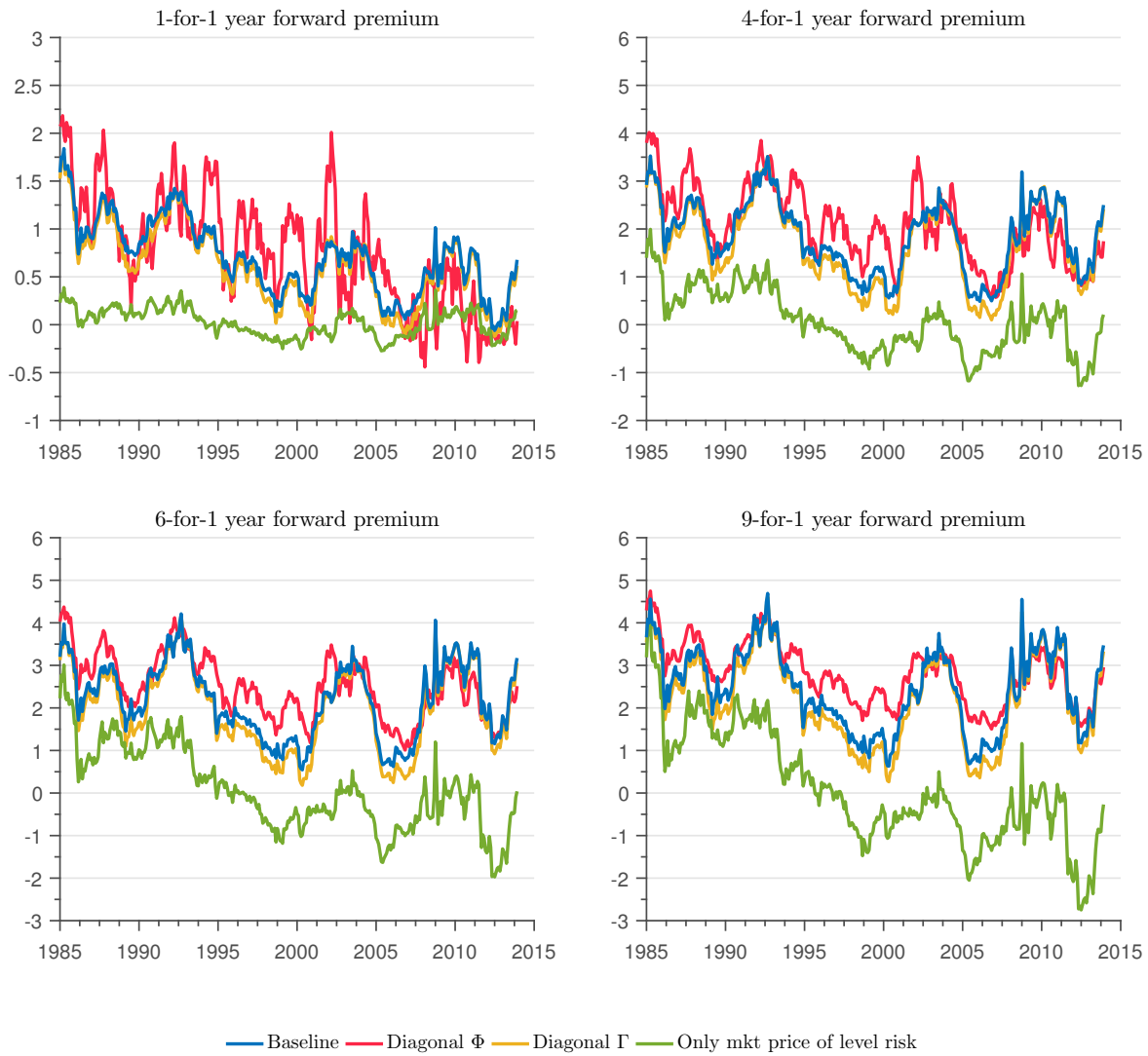


Figure 2. Forward premium: different models and bonds.

5.2. The Bond Holding Risk Premium

Given that the risk factors are virtually identical in the four models, differences in the estimated risk premiums come from differences in the factor loadings. Therefore, looking at factor loadings is a natural way to evaluate how sound the overidentifying restrictions are, we use Equation (16) to analyze those factor loadings in the case of the bond holding risk premium for different bond maturities and holding periods. Results are expressed in annualized percentage terms.

The three plots on the left panel of Figure 3 show the loadings on the factors $(\zeta_{1t}, \zeta_{2t}, \zeta_{3t})$, respectively, of the one-month bond holding risk premium as a function the maturity of the bond (buy and hold for one month an n period bond), for the four different models. The right panel displays the same loadings, but holding the position for one year.

The restrictions imposed on the models are economically relevant. In the baseline (unrestricted) model, and for both holding periods, the importance of the level factor ζ_{1t} increases with the maturity of the bond, while that of the slope factor ζ_{2t} is maximized (in absolute value) at about 84 months (seven years) when holding the bond for one month and is always decreasing in maturity when holding the bond for one year. The loading on the curvature factor ζ_{3t} is almost zero for bonds with maturities up to 40 months and then becomes negative and decreasing for bonds of longer maturities. In contrast, the model with a diagonal Φ matrix works quite differently: the contribution of the level

factor $\tilde{\zeta}_{1t}$ also increases (almost linearly) with the maturity of the bond, but by a much smaller rate than in the baseline model. On the other hand, the contribution of the curvature factor $\tilde{\zeta}_{3t}$ is now positive and increasing with the maturity of the bond for both holding periods. Those differences are economically relevant: a unit increase in the level factor leads to an increase in the one-month bond premium of over two percentage points for a 10-year bond in the baseline model, but of less than 0.5 percentage points in the diagonal Φ model.

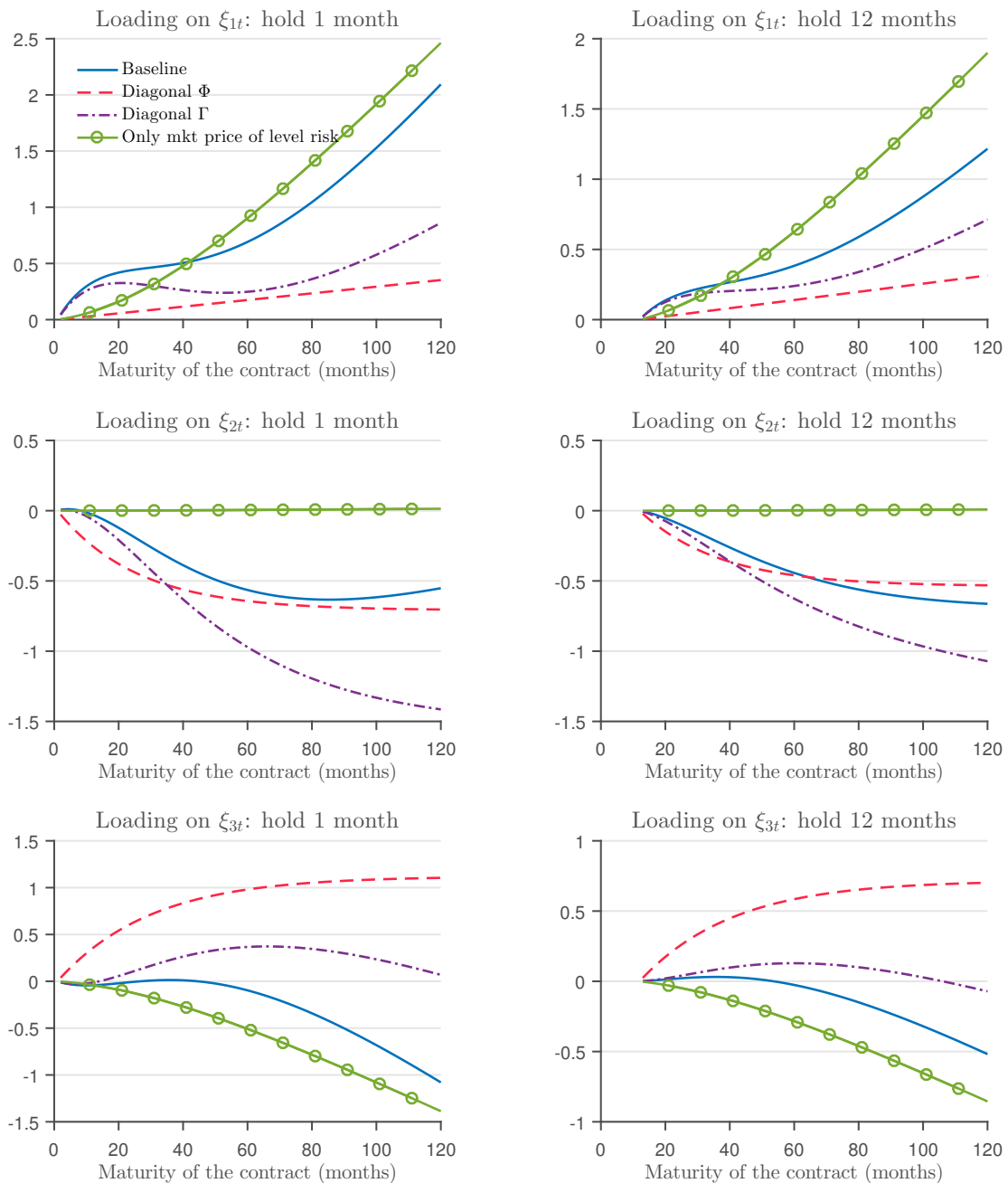


Figure 3. Loadings on factors: one-month and one-year excess holding returns.

Perhaps surprisingly, the factor loadings of the model with a diagonal Γ matrix are quite different from those of the baseline model. In the diagonal Γ model, the loading on the level factor is half of that in the baseline model, the loading on the slope factor is twice as large and the loading on the curvature factor is mostly positive, while that in the baseline model is mostly negative. Yet, as we

show below, those differences in the loadings tend to cancel out, and the evolution of the bond holding risk premium in the restricted model is similar to that in the baseline model.

The factor loadings of the model with only the market price of level risk (only the first row of λ_1 is non-zero) are very different from those of the baseline model. The influence of the slope factor is virtually zero for all bond maturities and holding periods. This prediction is at odds with those of the other models, in which the slope factor has a significant impact on all measures of risk premia. Interestingly, for this restricted model, the importance of the level factor when holding long-term bonds is not only very similar to that in the baseline model, but also dominates quantitatively the influence of the other factors. This observation suggests that the evolution of the bond holding risk premium will be very different from that of the baseline model when considering short maturity bonds, but similar when considering longer term bonds.

We next evaluate the evolution over time of the bond holding risk premium derived from the four models. If a restricted model is statistically rejected, but it generates an estimated risk premium that is similar to that in the baseline model, one could claim that the statistical rejection is not economically relevant.⁸ On the other hand, when a restricted model produces a risk premium that is substantially different from that in the baseline model, we say that the statistical rejection also has economic relevance.

Figure 4 shows one-month bond holding excess returns for 1-year, 5-year, 7-year and 10-year bonds for the different models. Consistent with the information in Figure 3, which shows that the loadings on the level and slope factors increase in absolute value with the maturity of the bond, we observe that the risk premia of long maturity bonds are substantially more volatile than those of shorter maturity bonds. The excess return from holding one-year bonds reaches a peak of around two percentage points, while those of the 5-, 7- and 10-year bonds are around six and 12 percentage points, respectively. Comparing this results with those presented in Figure 1, it is apparent that excess holding returns move closely with the slope factor, especially for longer maturity bonds.

The model with a diagonal Γ matrix is the one that displays the smaller differences in the bond holding risk premium relative to the baseline model: the differences are centered on zero, and their magnitudes are smaller than those of the other restricted models. Which model differs the most relative to the baseline, however, depends on the maturity of the bond. For the one-year bond, the model with only the market price of level risk is, by far, the most different from the baseline, with differences reaching 100 percentage points of excess holding returns. This model, however, delivers risk premiums similar to those of the baseline model for long-term bonds. As Figure 3 shows, in this model, the level factor dominates the contribution of the other two factors, and in that sense, it is similar to the baseline model. The second most different model is that with a diagonal Φ matrix. This model often produces differences of about 200 percentage points larger or smaller relative to the baseline (for example, in mid-2002).

⁸ Of course, those differences may be relevant in other dimensions such as forecasts of future yields. Nevertheless, models with good forecasting performance, such as the diagonal Φ model (Christensen et al. 2011), may produce poor estimates of risk premia in sample.

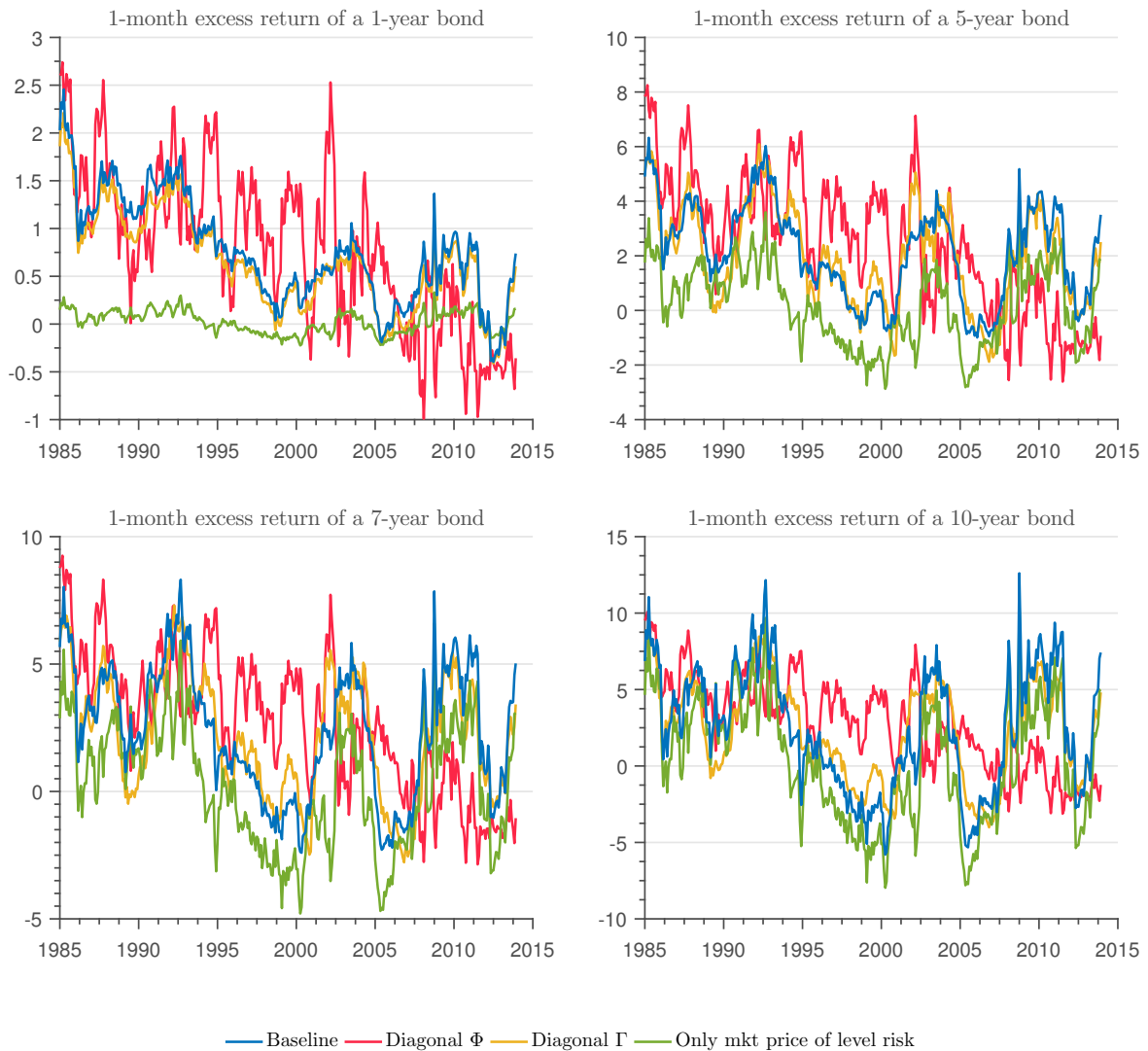


Figure 4. One-month expected excess returns: different models and bonds.

Figure 5 displays the 6-month, 1-year, 5-year and 7-year excess returns from holding a 10-year bond. A pattern that emerges from the figure is that the risk premiums associated with longer maturity bonds are highly contemporaneously correlated with the slope factor ζ_{2t} (see Figure 1). As with the one-month excess holding returns, the model that is closest to the baseline is that with a diagonal Γ matrix.

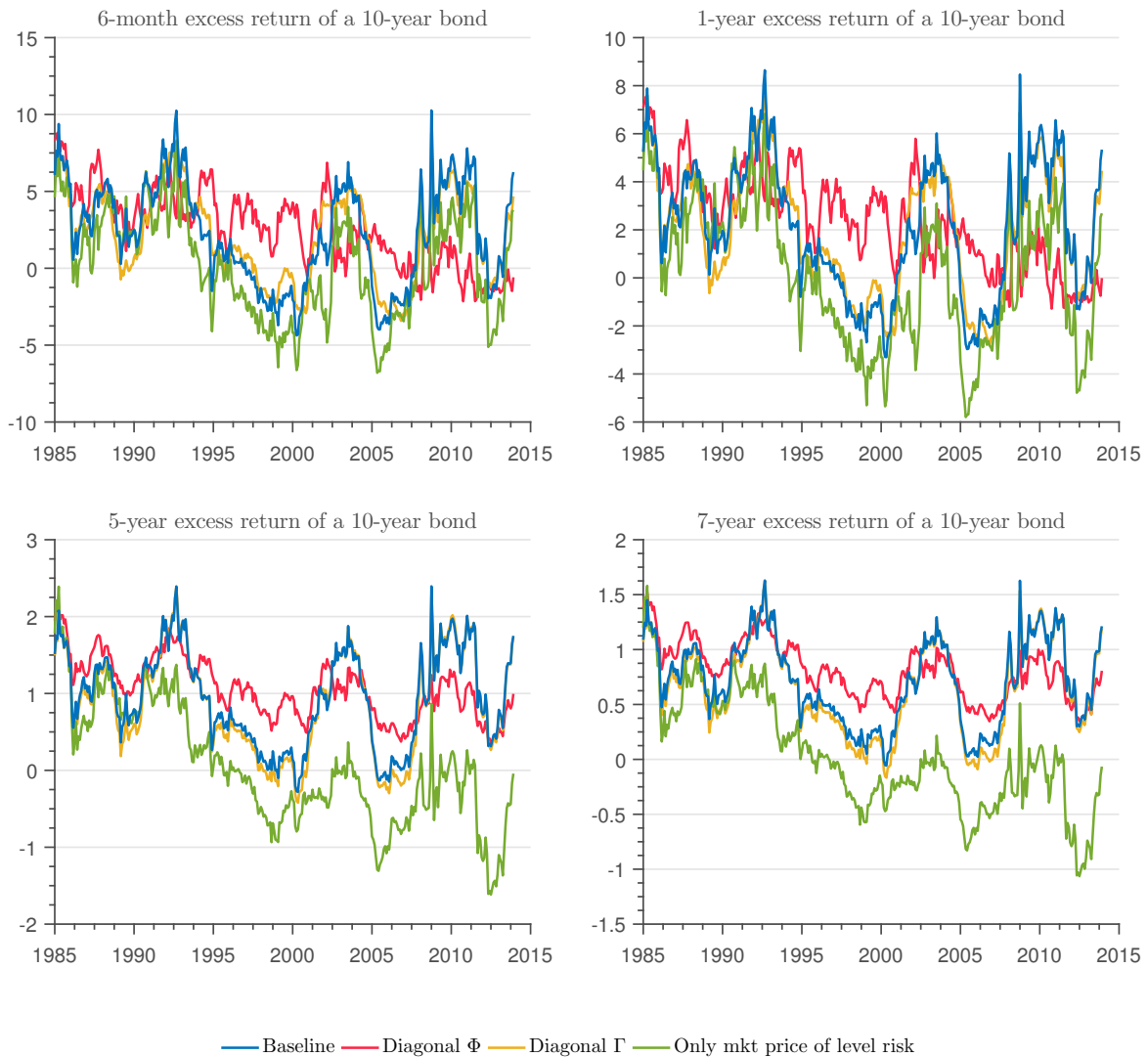


Figure 5. Expected returns on a 10-year bond: different models and holding periods.

5.3. The Term Premium

The n -period term premium is difference between the yield on an n -period zero coupon bond and the average of expected future short rates, where the average is taken between today and $n - 1$ periods ahead. Figure 6 shows, in the different panels, the evolution of the term premium using 1-year, 4-year, 6-year and 10-year bonds for the baseline and the restricted models. Two messages follow from this plot. First, the term premium moves closely with the slope factor (ξ_{2t}) especially for long maturity bonds. Second, the violations of the expectational hypothesis, as represented by the volatility of the term premium, become more apparent the longer the bond is under consideration. In particular, while the 10-year term premium can be as high as four percentage points and very volatile, the one-year term premium is smaller and less volatile, although it is far from constant and displays a decreasing trend over time. As for restrictions on risk prices, we still find that the model with a diagonal Γ matrix produces the term premium that is closest to the baseline.

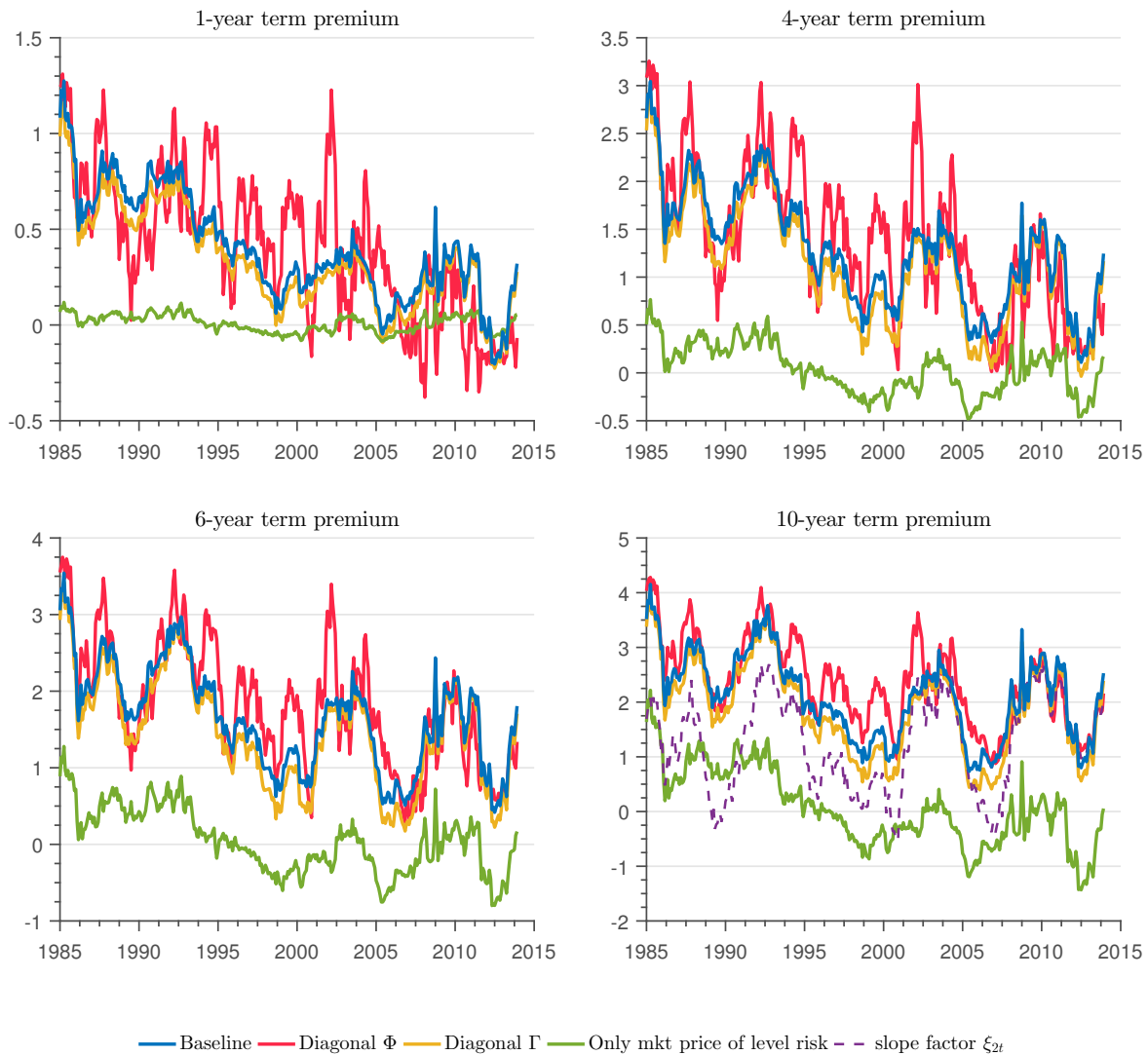


Figure 6. Term premium: different models and bonds.

5.4. The Holding Futures Risk Premium

Here, we consider the premium from holding a forward contract for a number of periods and selling it before the settlement date. Figure 7 displays the factor loadings of the one-month (left panels) and one-year (right panels) holding forward risk premium as a function of the settlement date, for the different models. As in the case of the bond holding return, the shape of the loadings differ substantially depending on the restrictions that we impose on the models. While the loadings on the level and slope factors (ζ_{1t} and ζ_{2t}) have a U-shape in the baseline model, the loading on the level factor in the model with a diagonal Γ matrix is flat and close to zero, and the loading on the slope factor is negative and increases towards zero in the model with a diagonal Φ matrix. Likewise, in the model with only market price of level risk, the loading on the level factor increases with the settlement date, and that on the slope factor is virtually zero.

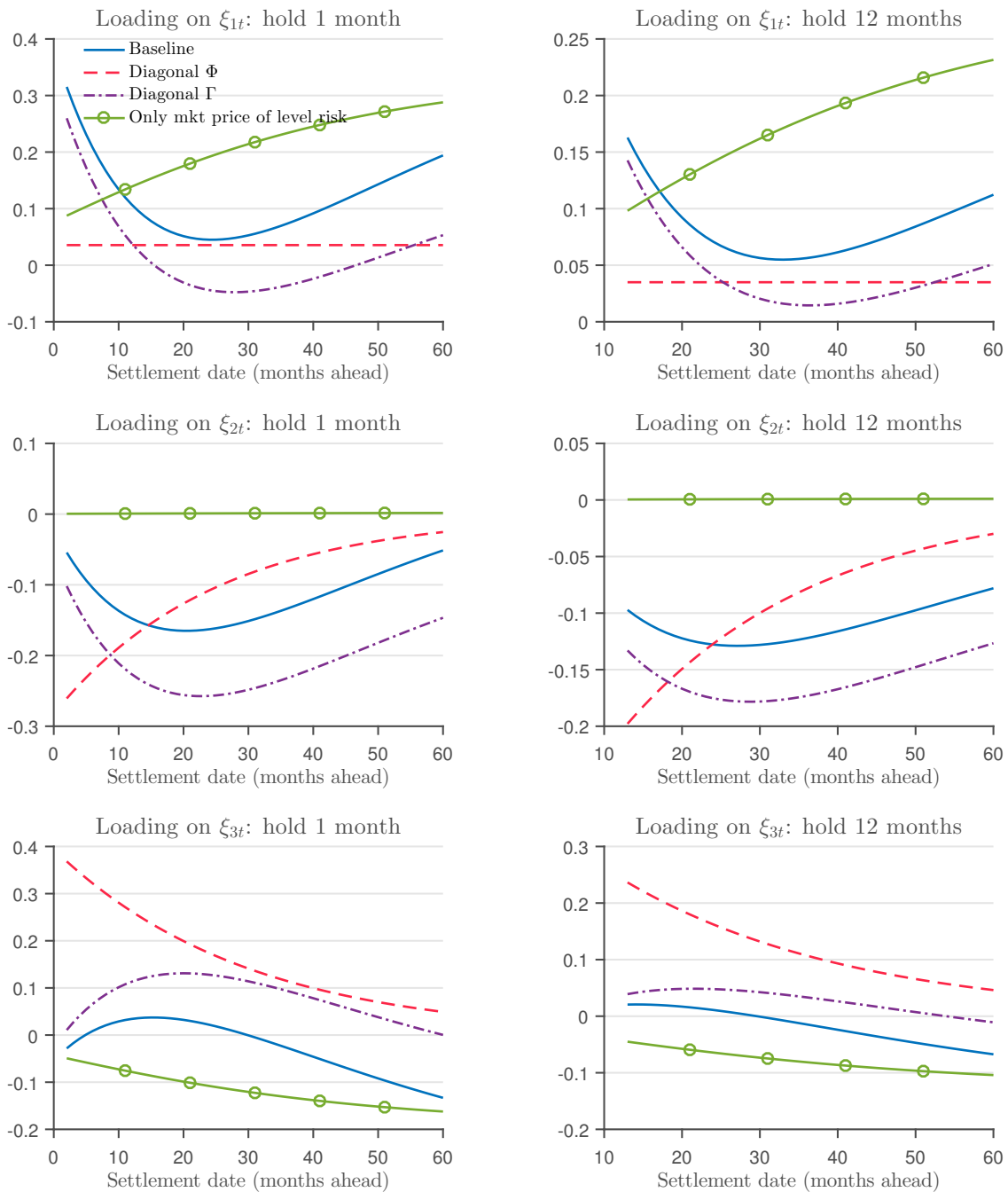


Figure 7. Loadings on factors: one-month and one-year holding forward returns.

Figures 8 and 9 show the evolution of the one-month and one-year holding forward risk premiums. Three patterns emerge from the figures. First, the holding forward risk premiums are much smaller than the bond holding risk premiums (excess bond returns). Second, the holding forward premium is also highly contemporaneously correlated with the slope factor ζ_{2t} . Third, the difference between the restricted models and the baseline tend to follow a pattern similar to the bond holding risk premium.

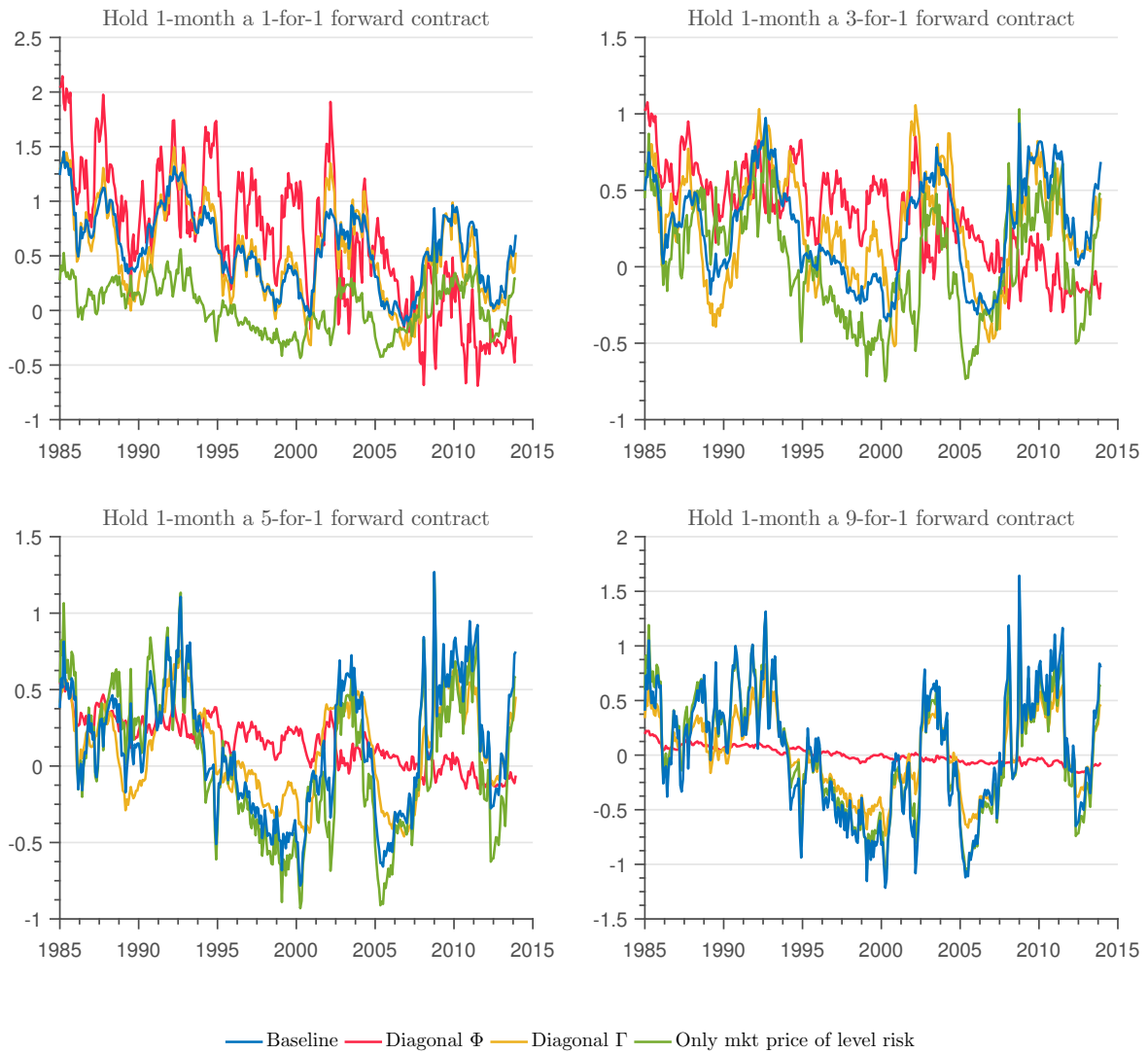


Figure 8. One-month holding forward risk premium.

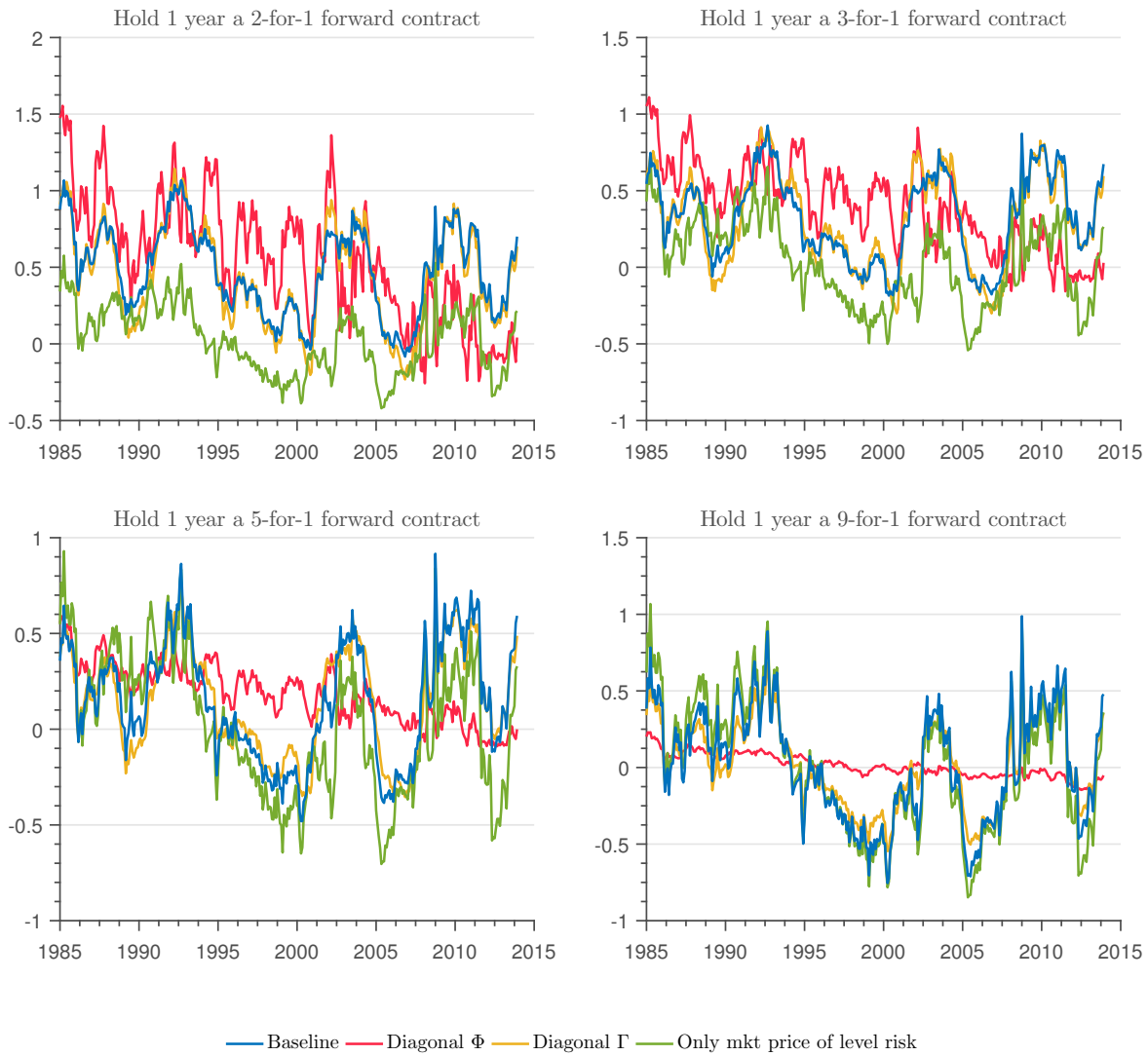


Figure 9. One-year holding forward risk premium.

6. Final Remarks

In this paper, we analyzed how different restrictions on risk prices affect bond risk premia. The model that imposes a diagonal Φ matrix and that with “only market price of level risk” are not only statistically rejected, but also produce measures of risk premia that are very different from those of the unrestricted model. Yet, there are models, such as that with a diagonal Γ matrix, that are also rejected on statistical grounds, but that produce measures of risk premia very similar to those from the baseline model. This result follows because restrictions on Γ have a minor impact on the market price of risk parameters. Any difference with the baseline model is of second order and comes from the indirect effect of the restrictions on the constants A_n , and because the estimates of the remaining, common, parameters may be somewhat different.

In the baseline model, the forward premium is always positive, while the excess holding risk premium is often negative. We found that the slope factor moves almost one-to-one with several measures of risk premia, which suggests that looking at the empirical slope of the yield curve (which is highly correlated with the slope factor that we extract from the model) gives an accurate description of risk premia in bond markets. Furthermore, as has been documented many times, we find that the expectational hypothesis of the term of the structure of interest rates is at odds with the data.

In terms of how restrictions on risk prices affect risk premia, we found that the differences across models in the estimated forward premium are modest, but they get amplified when considering excess holding risk premiums. There are, however, cases in which the restricted models seem to work well. For example, the one-month excess return of a 10-year bond for the model with “only market price of level risk” is quite similar to that from the baseline model. This happens because the loading of the risk premium on the level factor increases with the maturity of the bond and dominates movements in risk premia relative to the contribution of the other factors, but only for long maturity bonds.

In sum, imposing overidentifying restrictions affect the evolution of risk premia dramatically in some cases (such as when Φ is diagonal), but not so much in others (such as when Γ is diagonal). Often, researchers impose overidentifying restrictions to reduce the number of parameters to estimate or for other reasons, such as forecasting performance, even though they are statistically rejected. Our exercises suggest that this practice is inappropriate, since it is difficult to know a priori when the restrictions will have a significant impact on estimated risk premia. Of course, there is no harm in imposing the restrictions if they are not statistically rejected. However, since there is no theoretical guidance on how to impose them and one gains at best a few degrees of freedom by finding the appropriate restrictions, we believe that it is advisable to focus on the just-identified case whenever the objective of the study is to understand risk premia.

Author Contributions: C.H. and M.S. contributed equally to the paper.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

In the text, we claim that the four concepts of risk premia $\pi_{f,t}^{(s,n-s)}$, $\pi_{hb,t}^{(n,h)}$, $\pi_{y,t}^{(n)}$ and $\pi_{hf,t}^{(s,n-s,h)}$ are equivalent, in the sense that if one of them is zero or constant, so are the other three. That the risk premiums $\pi_{f,t}^{(s,n-s)}$, $\pi_{hb,t}^{(n,h)}$, and $\pi_{y,t}^{(n)}$ are equivalent is a well-known result. Here, we prove that the holding forward risk premium $\pi_{hf,t}^{(s,n-s,h)}$ is also equivalent to the other three.

Consider the equations that define the forward risk premium at times t and $t + h$,

$$\begin{aligned} f_t^{(s,n-s)} &= E_t \left[y_{t+s}^{(n-s)} \right] + \pi_{f,t}^{(s,n-s)}, \\ f_{t+h}^{(s-h,n-s)} &= E_{t+h} \left[y_{t+s}^{(n-s)} \right] + \pi_{f,t+h}^{(s-h,n-s)}. \end{aligned}$$

Taking expectations as of time t in the second equation gives:

$$E_t \left[f_{t+h}^{(s-h,n-s)} \right] = E_t \left[y_{t+s}^{(n-s)} \right] + E_t \left[\pi_{f,t+h}^{(s-h,n-s)} \right]$$

Subtracting this expression from the first equation, one obtains:

$$f_t^{(s,n-s)} - E_t \left[f_{t+h}^{(s-h,n-s)} \right] = \pi_{f,t}^{(s,n-s)} - E_t \left[\pi_{f,t+h}^{(s-h,n-s)} \right].$$

However, since the holding forward risk premium is given by:

$$\pi_{hf,t}^{(s,n-s,h)} = E_t \left[f_t^{(s,n-s)} - f_{t+h}^{(s-h,n-s)} \right],$$

it follows that:

$$\pi_{hf,t}^{(s,n-s,h)} = \pi_{f,t}^{(s,n-s)} - E_t \left[\pi_{f,t+h}^{(s-h,n-s)} \right].$$

Therefore, if the forward risk premium $\pi_{f,t}^{(s,n-s)}$ is zero (or constant), so is the holding forward risk premium $\pi_{hf,t}^{(s,n-s,h)}$ for any $(s, n - s, h)$.

Conversely, suppose that $\pi_{\text{hf},t}^{(s,n-s,h)}$ is zero or constant for all $(s, n - s, h)$. In particular, setting $s = h$ and $n - s = 1$ for any h , it follows that:

$$\begin{aligned} f_t^{(s,n-s)} &= E_t \left[f_{t+h}^{(s-h,n-s)} \right] + \pi_{\text{hf}}^{(s,n-s,h)}, \\ f_t^{(h,1)} &= E_t \left[f_{t+h}^{(0,1)} \right] + \pi_{\text{hf}}^{(h,1,h)}. \end{aligned}$$

However, $f_{t+h}^{(0,1)} = y_{t+h}^{(1)}$, so that:

$$f_t^{(h,1)} = E_t \left[y_{t+h}^{(1)} \right] + \pi_{\text{hf}}^{(h,1,h)}.$$

Rearranging and using the definition of $\pi_{f,t}^{(h,1)}$ gives $\pi_{f,t}^{(h,1)} = \pi_{\text{hf}}^{(h,1,h)}$. Since $\pi_{\text{hf}}^{(h,1,h)}$ is zero or constant, so is $\pi_{f,t}^{(h,1)}$ for all h . However, using $\pi_{y,t}^{(n)} = \frac{1}{n} \sum_{j=1}^n \pi_{f,t}^{(j,1)}$, it then follows that if $\pi_{\text{hf}}^{(h,1,h)}$ is zero or constant, so is $\pi_{y,t}^{(n)}$ for all n . Therefore, the holding forward risk premium is equivalent to the term premium and, therefore, to the other two definitions of risk premium. This completes the proof that the four definitions of risk premiums are equivalent.

References

- Ang, Andrew, and Monika Piazzesi. 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50: 745–87. [CrossRef]
- Bauer, Michael D. 2018. Restrictions on risk prices in dynamic term structure models. *Journal of Business & Economic Statistics* 36: 196–211.
- Campbell, John, and Robert Shiller. 1991. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies* 58: 495–514. [CrossRef]
- Christensen, Jens H. E., and Glenn D. Rudebusch. 2015. Estimating shadow-rate term structure models with near-zero yields. *Journal of Financial Econometrics* 13: 226–59. [CrossRef]
- Christensen, Jens, Jose Lopez, and Glenn Rudebusch. 2010. Inflation expectations and risk premiums in an arbitrage-free model of nominal and real bond yields. *Journal of Money, Credit and Banking* 42: 143–78. [CrossRef]
- 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* 164: 4–20. [CrossRef]
- Cochrane, John H. 2015. *Comments on "Robust Bond Risk Premia" by Michael Bauer and Jim Hamilton*. Working paper. Chicago: University of Chicago, unpublished.
- Cochrane, John H., and Monika Piazzesi. 2005. Bond risk premia. *American Economic Review* 95: 138–60. [CrossRef]
- Cochrane, John H., and Monika Piazzesi. 2008. *Decomposing the Yield Curve*. Technical Report 18. Society for Economic Dynamics.
- Dai, Qiang, and Kenneth J. Singleton. 2000. Specification analysis of affine term structure models. *Journal of Finance* 55: 1943–78. [CrossRef]
- Diebold, Francis X., and Canlin Li. 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130: 337–64. [CrossRef]
- Diebold, Francis X., and Glenn D. Rudebusch. 2012. *Yield Curve Modeling and Forecasting: The Dynamic Nelson–Siegel Approach*. Number 9895 in Economics Books. Princeton: Princeton University Press.
- Duffee, Gregory R. 2011. *Forecasting with the Term Structure: The Role of No-Arbitrage Restrictions*. Economics Working Paper Archive 576. Baltimore: The Johns Hopkins University, Department of Economics.
- Evgenidis, Anastasios, Athanasios Tsagkanos, and Costas Siriopoulos. 2017. Towards an asymmetric long run equilibrium between stock market uncertainty and the yield spread. A threshold vector error correction approach. *Research in International Business and Finance* 39: 267–79. [CrossRef]
- Fama, Eugene, and Robert R. Bliss. 1987. The information in long-maturity forward rates. *American Economic Review* 77: 680–92.
- Gürkaynak, Refet S., and Jonathan H. Wright. 2012. Macroeconomics and the Term Structure. *Journal of Economic Literature* 50: 331–67. [CrossRef]

- Hamilton, James D., and Jing Cynthia Wu. 2012. Identification and estimation of gaussian affine term structure models. *Journal of Econometrics* 168: 315–31. [[CrossRef](#)]
- Hevia, Constantino, Martin Gonzalez Rozada, Martin Sola, and Fabio Spagnolo. 2015. Estimating and Forecasting the Yield Curve Using A Markov Switching Dynamic Nelson and Siegel Model. *Journal of Applied Econometrics* 30: 987–1009. [[CrossRef](#)]
- Joslin, Scott, Kenneth Singleton, and Haoxiang Zhu. 2011. A new perspective on gaussian dynamic term structure models. *Review of Financial Studies* 24: 926–70. [[CrossRef](#)]
- Nelson, Charles R., and Andrew F. Siegel. 1987. Parsimonious Modeling of Yield Curves. *The Journal of Business* 60: 473–89.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).