An asset pricing approach to testing general term structure models

Online Appendix

Bent Jesper Christensen[a,b,c], Michel van der Wel[b,d,e,f]

[a] *Aarhus University, Department of Economics and Business Economics, Fuglesangs Alle 4, 8210 Aarhus V, Denmark*
[b] *Center for Research in Econometric Analysis of Time Series, Aarhus University, Department of Economics and Business Economics, Fuglesangs Alle 4, 8210 Aarhus V, Denmark*
[c] *Dale T. Mortensen Center, Aarhus University, Department of Economics and Business Economics, Fuglesangs Alle 4, 8210 Aarhus V, Denmark*
[d] *Erasmus University Rotterdam, Erasmus School of Economics, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands*
[e] *Tinbergen Institute, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands*
[f] *Erasmus Research Institute of Management, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands*

Abstract

This is the online appendix for "An asset pricing approach to testing general term structure models" by Bent Jesper Christensen and Michel van der Wel.

Contents

Appendix A. Additional Tables and Figures

Table OA1
Estimated volatility functions.

This table reports the estimated volatility function for the $d = 4$ factor model with time-varying λ_t where the no-arbitrage restriction is imposed. Asymptotic t-statistics are given below the estimates in parentheses, and * and ** denote significance at the 5% and 1% level, respectively.

Volatility function estimates $d = 4$				
	Factor 1	Factor 2	Factor 3	Factor 4
$b_{1,j}$	0.0206^{**} (15.2)	0	0	0
$b_{2,j}$	0.0155^{**} (5.09)	0.00321 (1.05)	0	0
$b_{3,j}$	0.0132^{**} (9.88)	0.00323^{*} (2.43)	0.00283^{*} (2.15)	0
$b_{4,j}$	0.0119^{**} (4.72)	0.00388 (1.54)	0.00473 (1.89)	0.0000000225 (0.00000939)
$b_{5,j}$	0.0104^{**} (40.6)	0.00449^{**} (17.8)	0.00382^{**} (15.4)	0.00173^{**} (20.7)
$b_{6,j}$	0.00997^{**} (48.8)	0.00387^{**} (19.2)	0.00415^{**} (20.9)	0.00259^{**} (38.7)
$b_{7,j}$	0.00926^{**} (37.1)	0.00339^{**} (13.8)	0.00395^{**} (16.3)	0.00324^{**} (39.7)
$b_{8,j}$	0.00885^{**} (67.5)	0.00244^{**} (18.9)	0.00557^{**} (43.8)	0.00324^{**} (75.2)
$b_{9,j}$	0.00886^{**} (123)	0.00118^{**} (16.6)	0.00593^{**} (84.9)	0.00428^{**} (180)
$b_{10,j}$	0.00865^{**} (126)	-0.000242^{**} (−3.58)	0.00699^{**} (105)	0.00458^{**} (201)
$b_{11,j}$	0.00822^{**} (140)	-0.00177^{**} (−30.7)	0.00659^{**} (116)	0.00495^{**} (254)
$b_{12,j}$	0.00769^{**} (211)	-0.00238^{**} (−65.9)	0.00715^{**} (203)	0.00452^{**} (368)
$b_{13,j}$	0.0077^{**} (143)	-0.00296^{**} (−55.7)	0.00632^{**} (121)	0.00477^{**} 266)
$b_{14,j}$	0.00704^{**} (159)	-0.00317^{**} (−72.5)	0.00613^{**} (143)	0.00424^{**} (286)
$b_{15,j}$	0.00679^{**} (178)	-0.00365^{**} (−96.6)	0.00618^{**} (167)	0.00414^{**} (323)
$b_{16,j}$	0.00658^{**} (115)	-0.00413^{**} (−73.1)	0.00562^{**} (101)	0.00436^{**} (228)

Table OA2

Regression of risk prices on macroeconomic variables.

In this table we regress the time-varying risk prices from our model on the macroeconomic variables that we consider. The risk premium estimates used for constructing the correlation matrix correspond to the restricted models with time-varying risk premiums. The macroeconomic variables we consider are industrial production (“Ind Prod”), nonfarm payroll employment (“Empl”), and the second and fourth principal components (“PC 2” and “PC 4”) of the Stock and Watson data set of 111 variables without missing observations from January 1985 through December 2016 from the McCracken and Ng (2016) FRED-MD data set. We show two variants of each regression for each factor. We first regress the relevant risk price on the individual macro variable, then condition on the macro variables for previous risk factors in the regression. Robust Newey and West (1987) t-statistics are given below the estimates in parentheses, and * and ** denote significance at the 5% and 1% level, respectively. In addition we report the adjusted R^2 and the number of observations.

	Risk price regressions						
	f1 (from $d = 1$)	f2 (2nd from $d = 2$)		f3 (2nd from $d = 3$)		f4 (4th from $d = 4$)	
	(1)&(2)	(1)	(2)	(1)	(2)	(1)	(2)
PC 2	-0.00919^{**}		-0.0164^{**}		0.00247		-0.0505^{**}
	(−4.56)		(−4.30)		(0.373)		(−2.69)
Ind Prod		0.0554^{**}	0.0782^{**}		-0.0999^{*}		0.00611
		(3.03)	(3.82)		(−1.99)		(0.0656)
PC 4				0.0408^{**}	0.0436^{**}		0.00682
				(3.50)	(3.73)		(0.268)
Empl						0.00228^{**}	0.00222^{**}
						(6.22)	(5.42)
Intercept	-0.380^{**}	-0.0786^{**}	-0.0821^{**}	-0.329^{**}	-0.313^{**}	0.0446	0.0514
	(−32.5)	(−4.67)	(−5.03)	(−8.24)	(−7.72)	(0.462)	(0.537)
R^2	0.037	0.026	0.084	0.052	0.070	0.225	0.254
#Obs	382	382	382	382	382	382	382

Table OA3
Estimated volatility functions, including macro factors.

This table reports the estimated volatility functions for the $d = 4$ factor model with $d^o = 2$ macro factors in which λ_t is time-varying and the no-arbitrage restriction imposed. The left part shows the volatility functions for the unobserved factors, and the right part for the observed macroeconomic factors. The macro factors in this model are the second principal component from a large set of macroeconomic variables, and industrial production. Asymptotic t-statistics are given below the estimates in parentheses, and * and ** denote significance at the 5% and 1% level, respectively.

Volatility function estimates $(d, d^o) = (4, 2)$				
	B^u		B^o	
	Latent Factor 1	Latent Factor 2	Macro Factor 1	Macro Factor 2
$b_{1,j}$	0.0126** (44.6)	0	−0.000503** (−4.87)	0.000492** (4.94)
$b_{2,j}$	0.0123** (18.5)	0.00313** (2.67)	−0.000173** (−3.49)	0.000173** (3.59)
$b_{3,j}$	0.0104** (18.5)	0.00283** (2.71)	−0.00019** (−4.72)	0.000191** (4.88)
$b_{4,j}$	0.0106** (19.2)	0.00487** (4.03)	0.000311** (8.14)	−0.00029** (−7.76)
$b_{5,j}$	0.00953** (20.4)	0.00718** (11.4)	−0.0000621** (−4.95)	0.0000692** (5.72)
$b_{6,j}$	0.00853** (18.8)	0.00818** (14.6)	−0.000263** (−25.3)	0.000263** (26.3)
$b_{7,j}$	0.00784** (16.2)	0.00961** (16.7)	−0.00038** (−22.9)	0.000376** (23.4)
$b_{8,j}$	0.00667** (15.3)	0.00941** (18.6)	−0.000298** (−30.7)	0.000298** (31.7)
$b_{9,j}$	0.00537** (12.4)	0.0105** (24.6)	−0.000534** (−82)	0.000527** (84.4)
$b_{10,j}$	0.00441** (10)	0.0115** (25.8)	−0.000475** (−63.2)	0.000471** (64.8)
$b_{11,j}$	0.00293** (7.11)	0.0115** (29.5)	−0.000552** (−104)	0.000544** (107)
$b_{12,j}$	0.00253** (6.37)	0.0111** (26.8)	−0.000403** (−76.9)	0.000401** (78.9)
$b_{13,j}$	0.00181** (4.88)	0.0107** (31.1)	−0.000546** (−114)	0.000538** (116)
$b_{14,j}$	0.00169** (4.75)	0.0102** (27.2)	−0.000377** (−78.3)	0.000375** (80.1)
$b_{15,j}$	0.00134** (3.92)	0.00989** (28.7)	−0.000411** (−85.9)	0.000407** (87.6)
$b_{16,j}$	0.000904** (2.65)	0.00986** (31)	−0.000521** (−76.8)	0.000512** (77.9)

Table OA4
Sensitivity analysis of model performance overview, including macro factors.

This table provides two sensitivity analyses of the model performance measures. Panel A reports the performance measures for the case in which the nonlinear terms is omitted and raw yield changes are used instead of slope-adjusted yield changes. Panel B reports the measures when restricting B to affine subclasses. In all cases, models are estimated with varying number of covariance-generating factors $d = 1, 2, 3, 4$ and varying number of macro factors $d^o = 0, 1, \ldots, d$. In the case with the restricted B the ordering is such that the first factors are macro factors (if any are included) and the last (if any) are the unobserved factors. The table focusses on the case with time-varying risk premiums, and considers both an unrestricted version ("Unr") and a model imposing the no-arbitrage drift restriction ("Restr"). The table reports the log likelihood ("Likelihood") and the number of parameters for each of the models and provides the Akaike and Schwarz Bayesian information criteria ("AIC" and "BIC"), with the best value for each model highlighted in boldface. For the d^o=3 case, as the loadings on the 2nd and 3rd macro factors are similar we restrict $a_3 = 0$ and $A_{31} = A_{32} = 0$.

Panel A: Likelihood and number of parameters – APT on yield changes								
	Number of risk factors							
Number of	$d = 1$		$d = 2$		$d = 3$		$d = 4$	
macro factors	Unr	Restr	Unr	Restr	Unr	Restr	Unr	Restr
d^o=0								
Likelihood	33,880	33,879	36,856	36,854	37,619	37,619	37,834	37,833
No. of Pars	49	34	67	53	86	73	106	94
AIC	-67,662	**-67,690**	-73,578	**-73,602**	-75,066	**-75,091**	-75,455	**-75,478**
BIC	-67,332	**-67,461**	-73,127	**-73,246**	-74,488	**-74,601**	-74,743	**-74,847**
d^o=1								
Likelihood	27,172	27,171	34,696	34,695	37,623	37,622	38,381	38,381
No. of Pars	52	37	69	55	87	74	106	94
AIC	-54,241	**-54,269**	-69,254	**-69,280**	-75,072	**-75,096**	-76,551	**-76,574**
BIC	-53,888	**-54,018**	-68,782	**-68,904**	-74,477	**-74,590**	-75,826	**-75,931**
d^o=2								
Likelihood			27,092	27,092	35,605	35,605	38,417	38,417
No. of Pars			76	62	93	80	111	99
AIC			-54,033	**-54,061**	-71,024	**-71,049**	-76,613	**-76,636**
BIC			-53,513	**-53,637**	-70,378	**-70,494**	-75,842	**-75,949**
d^o=3								
Likelihood					28,030	28,030	36,512	36,512
No. of Pars					102	88	119	106
AIC					-55,857	**-55,885**	-72,787	**-72,812**
BIC					-55,154	**-55,278**	-71,949	**-72,066**
d^o=4								
Likelihood							28,146	28,145
No. of Pars							133	120
AIC							-56,026	**-56,050**
BIC							-55,102	**-55,217**

Panel B: Likelihood and number of parameters – Restricted B								
	Number of risk factors							
Number of	$d=1$		$d=2$		$d=3$		$d=4$	
macro factors	Unr	Restr	Unr	Restr	Unr	Restr	Unr	Restr
d^o=0								
Likelihood	27,102	26,809	29,147	28,925	29,877	29,634	29,995	29,874
No. of Pars	34	19	39	25	46	33	55	43
AIC	**-54,137**	-53,580	**-58,216**	-57,800	**-59,662**	-59,202	**-59,880**	-59,662
BIC	**-53,908**	-53,453	**-57,953**	-57,632	**-59,353**	-58,980	**-59,511**	-59,373
d^o=1								
Likelihood	23,122	20,932	28,149	27,945	28,134	27,993	30,571	30,361
No. of Pars	37	22	40	26	45	32	52	40
AIC	**-46,169**	-41,820	**-56,219**	-55,839	**-56,178**	-55,922	**-61,038**	-60,641
BIC	**-45,918**	-41,671	**-55,945**	-55,661	**-55,870**	-55,703	**-60,682**	-60,368
d^o=2								
Likelihood			23,093	23,067	28876	28,569	30,745	30,588
No. of Pars			47	33	50	37	55	43
AIC			**-46,093**	-46,067	**-57,653**	-57,063	**-61,380**	-61,091
BIC			-45,771	**-45,841**	**-57,306**	-56,806	**-60,998**	-60,792
d^o=3								
Likelihood					22,990	22,962	29,738	29,640
No. of Pars					59	45	62	49
AIC					**-45,862**	-45,834	**-59,352**	-59,182
BIC					-45,455	**-45,524**	**-58,916**	-58,837
d^o=4								
Likelihood							23,041	23,005
No. of Pars							76	63
AIC							**-45,931**	-45,885
BIC							-45,403	**-45,447**

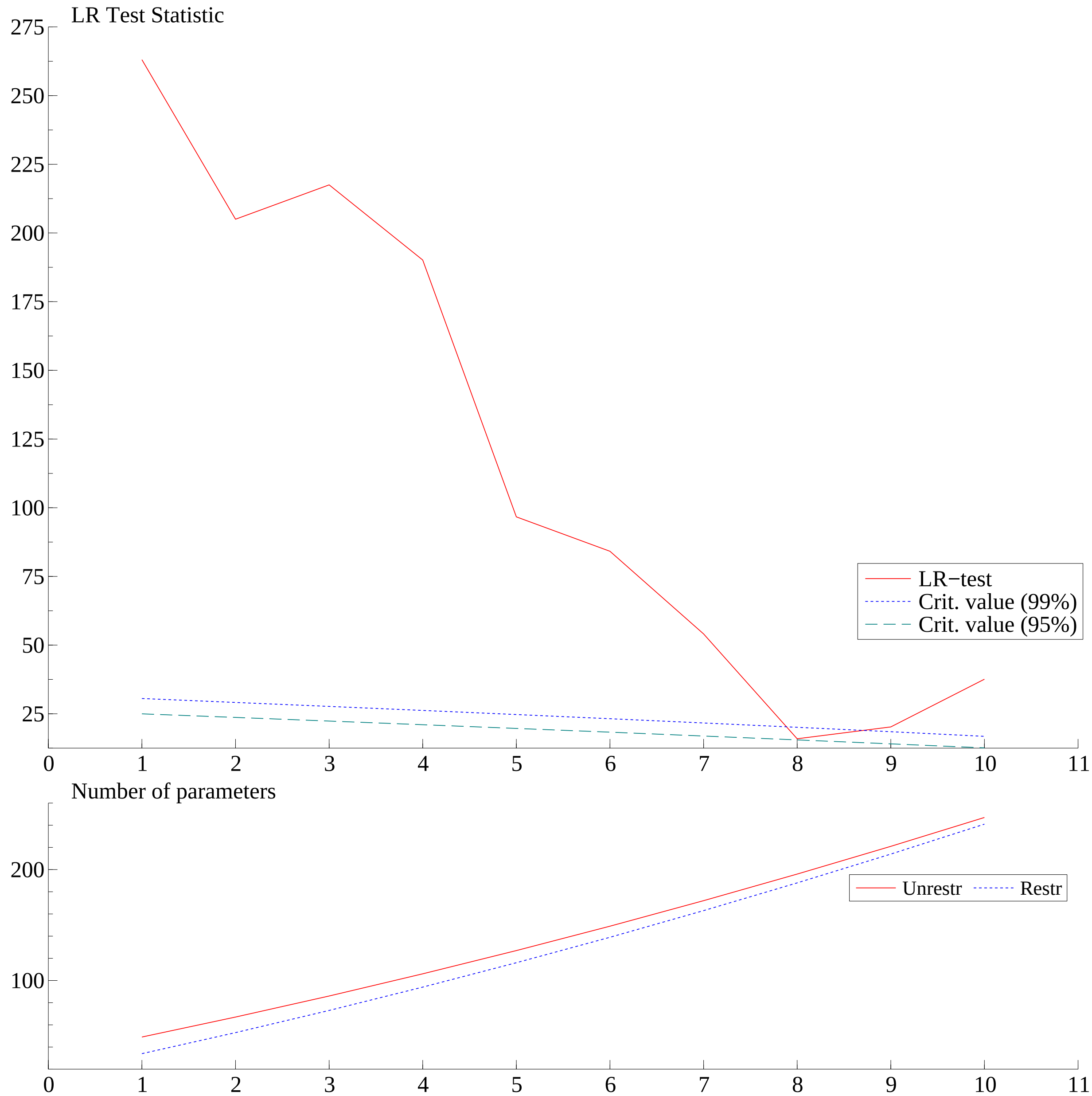


Fig. OA1. This figure shows results of tests of the no-arbitrage drift restriction in models with time-varying risk prices λ_t. For each number of covariance-generating factors $d = 1, 2, \ldots, 10$, models are estimated with and without the no-arbitrage drift restriction imposed. The top plot provides the LR test statistics, together with boundaries for 5% and 1% critical values. The bottom plot provides the number of parameters.

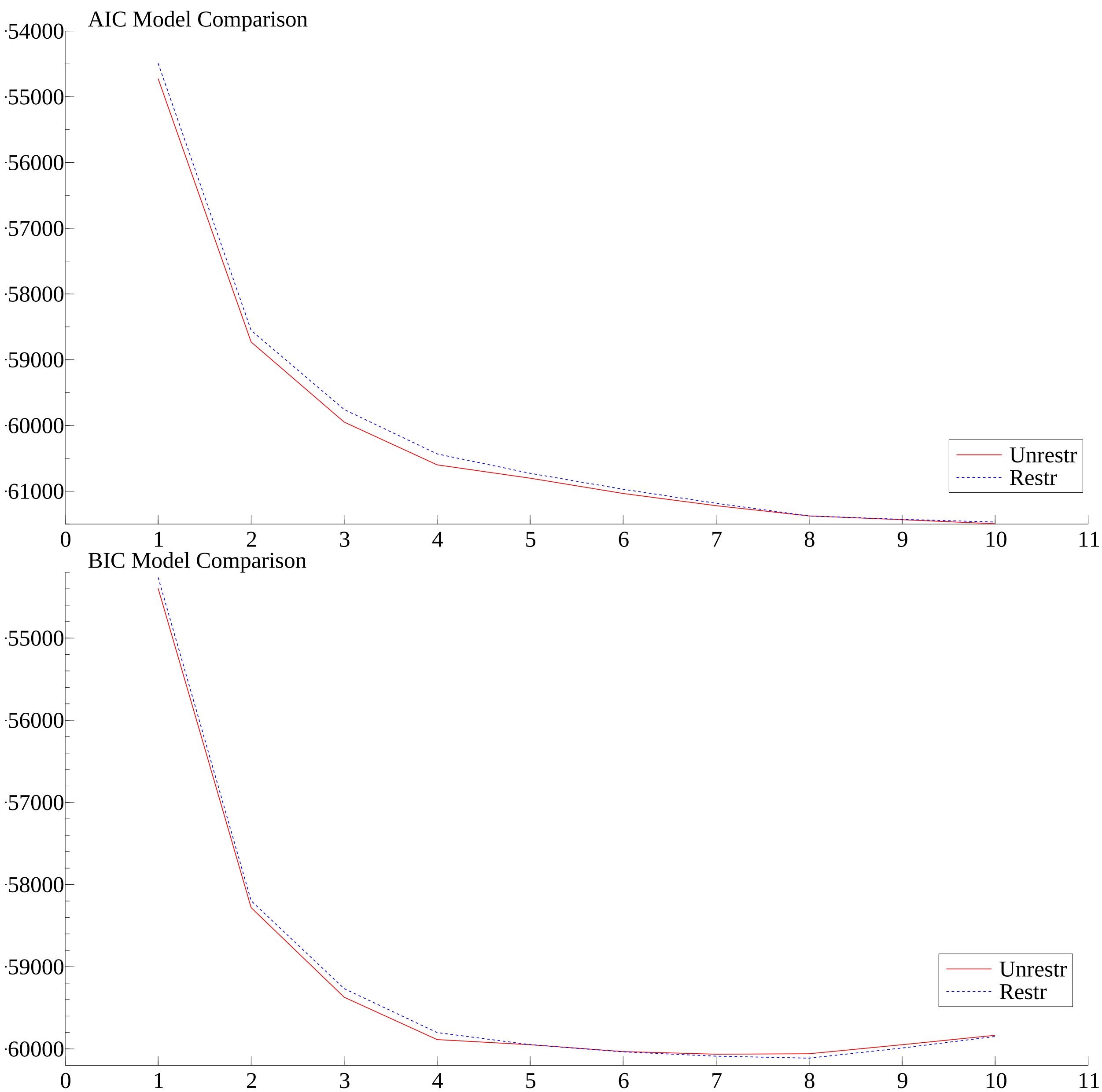


Fig. OA2. This figure shows performance based on information criteria for models with time-varying risk prices λ_t. For each number of covariance-generating factors $d = 1, 2, \ldots, 10$, models are estimated with and without the no-arbitrage drift restriction imposed. The top plot provides the AIC of models, and the bottom plot the BIC.

Panel A: Time-varying test probabilities

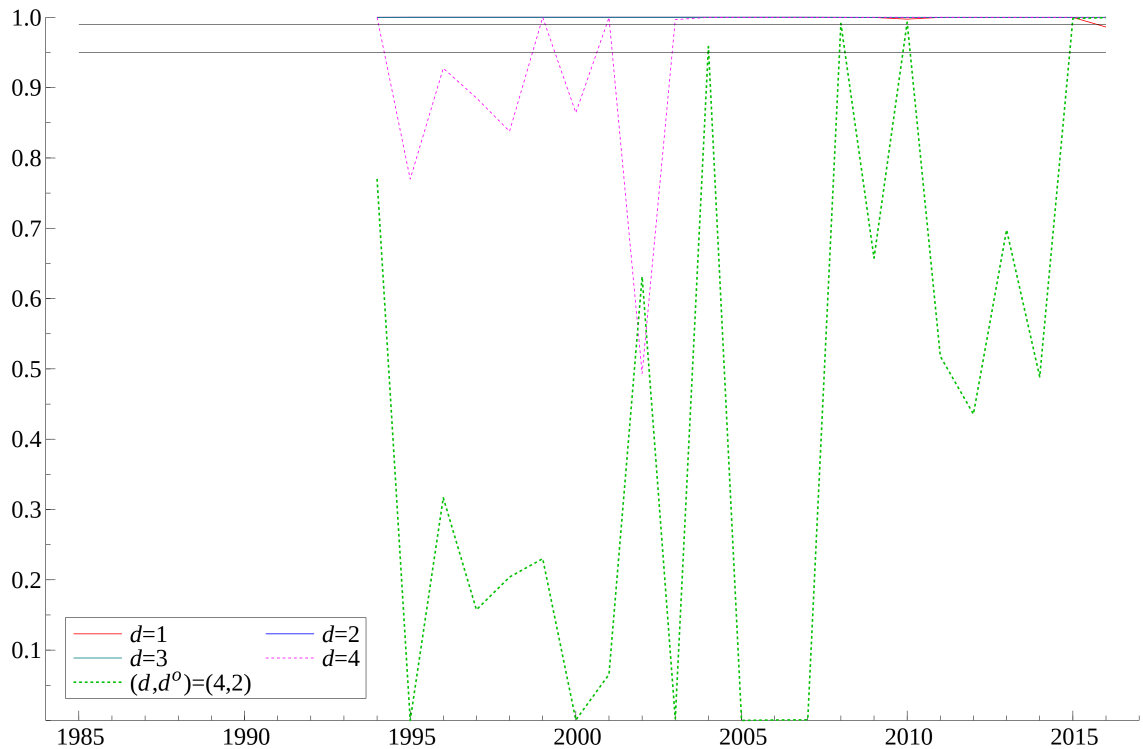
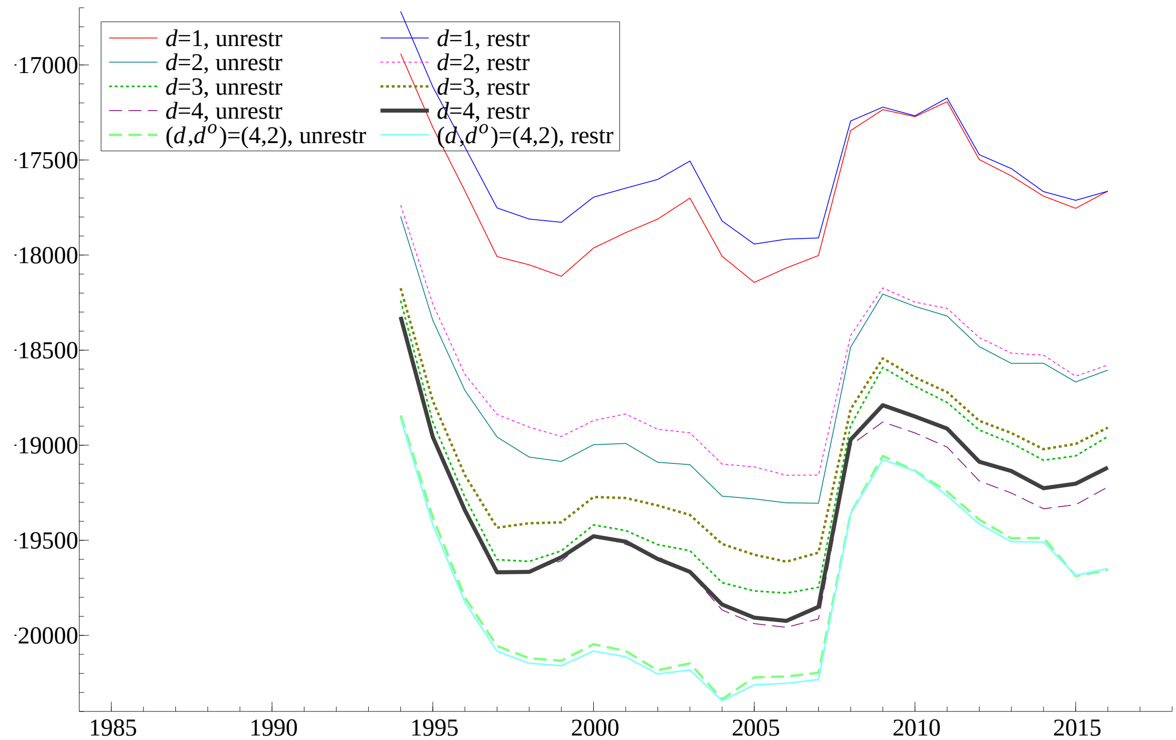
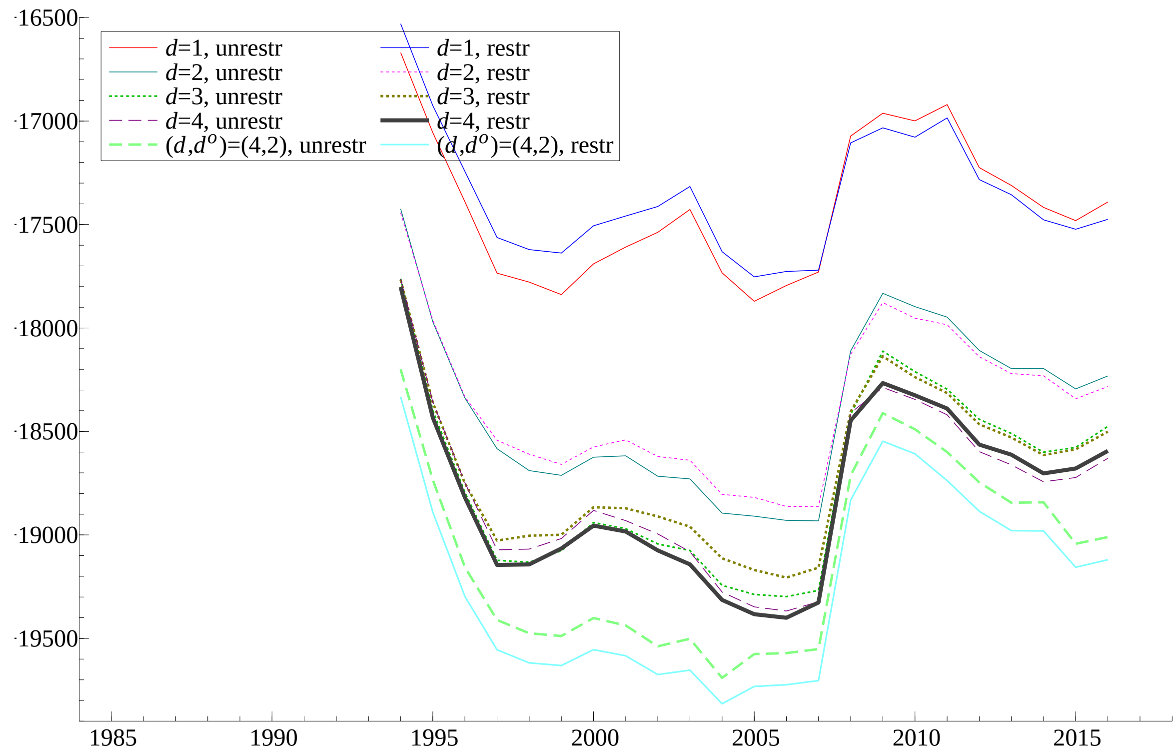


Fig. OA3. This figure shows LR tests of the no-arbitrage drift restriction over rolling ten-year windows. Models are estimated with varying number of covariance-generating factors $d = 1, 2, 3, 4$ and no macro factors, $d^o = 0$, in addition to a model with $(d, d^o) = (4, 2)$. The focus is on the case with time-varying risk premiums, both unrestricted and imposing the no-arbitrage drift restriction. Panel A depicts the p-value of rejecting the null hypothesis of the restricted case which imposes no-arbitrage. The horizontal lines depict 95% and 99%, so denote significant rejection of the no-arbitrage restriction at the 5% and 1% level, respectively. Panels B and C provide the AIC and BIC values, respectively.

Panel B: Time-varying AIC values



Panel C: Time-varying BIC values



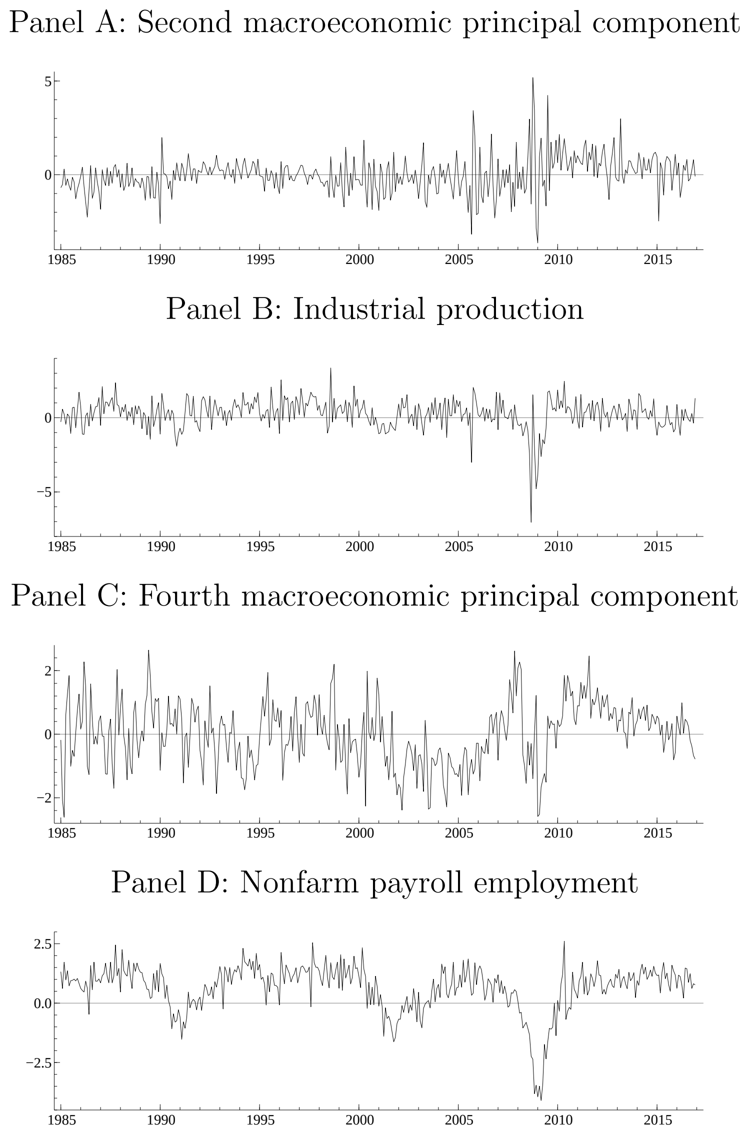


Fig. OA4. This figure shows time series plots of the macroeconomic variables in our analysis. These are the second and fourth principal components from a broad set of macroeconomic variables, industrial production (IP), and nonfarm payroll employment. For IP and employment, monthly relative changes are considered. The two principal components are obtained by running a principal component analysis on the Stock and Watson data set of 111 variables without missing observations from January 1985 through December 2016 from the McCracken and Ng (2016) FRED-MD data set.

Appendix B. Sensitivity to distributional assumptions

To see how robust our results are to alternative distributional assumptions we also estimate two model variations in which we use principal components analysis (PCA). In a first step we run a principal component analysis on the demeaned slope-adjusted yield changes. In particular, writing the PCA as $\tilde{y}_t - \hat{\mu} = Bw_t + \varepsilon_t$, where $\hat{\mu}$ denotes the sample mean, we also obtain estimates for the volatility matrix B and the covariance generating factors w_t in this framework. For comparison with the state space model, we use a version of the PCA with $var(w_t) = I_d$, absorbing the factor variances and covariances into B.

In Panel (A) of Fig. B1 we show the estimated loadings on the covariance generating factors. The pattern is very similar to that in Panel (A) of Fig. 2. The first factor is a slope factor, the second one, too, but with more curvature, even a hump around $i = 10$ (4 years to maturity). The third factor captures the more distinct hump at 15 months. In general, the PCA loadings on the first k factors are unchanged as more factors are added, a property not shared with the previous state space analysis.

In Panel (B) of Fig. B1 we provide a time series plot of the fitted covariance-generating factors $\widehat{w}_t$. Again, the picture is similar to that in Panel (B) of Fig. 2, although the factor scores exhibit stronger patterns across time and appear less similar to true noise in the PCA case, compared to the state space case.

The estimated volatility functions B and covariance generating factors w_t allow taking another look at the obtained risk prices. To this end, the regression of the sample estimate of the unconditional mean of the slope-adjusted yield changes $\hat{\mu}$ on the estimated volatility matrix from the principal component analysis $\hat{B}$ produces estimated risk prices. Following Eq. (6), we need to subtract $\hat{b}_i'\hat{b}_i\tau_i/2$ from each cross-sectional observation. The regression in this case is

$$\hat{\mu} - \mathrm{vec}_{1:m}\{\hat{b}_i'\hat{b}_i\tau_i/2\} = \hat{B}\lambda + \eta. \qquad \text{(Appendix B.1)}$$

An important thing to note is that this procedure is only an approximate way to get the risk prices. The estimates of the volatility functions are obtained in the first step through principal components. Thus, the information from the no-arbitrage relation estimated in the second step is not utilized when the volatility functions and covariance generating factors are obtained in the first. The state space approach allows imposing the drift restriction from the outset.

In Table B1 we show the estimated risk prices when we estimate Eq. (Appendix B.1) by generalized least squares (GLS), to account for heterogeneity. The covariance used to weight the observations is estimated using the PCA results: $\widehat{var[\eta]} = \hat{B}\hat{B}' + \hat{\Psi}$, where $\hat{\Psi}$ is diagonal with elements $\hat{\Psi}_{ii} = (1/T)\sum_t \hat{\varepsilon}_{it}^2$ for $i = 1, \ldots, m$. The risk prices we obtain follow a similar pattern to those in the state space model in Table 2, keeping in mind that the ordering of the factors in the table switches in the state space case. In both cases, the leading risk price is largest in magnitude, and all prices are negative, except that on the third factor in the PCA case (fourth in the state space model). The main difference is that the risk prices in the PCA are all insignificant, suggesting less precision in this approach, compared to the state space model.

To obtain time-varying estimates of the risk prices in the PCA framework we estimate a slightly altered version of the above. Based on Eqs. (12) and (9), the relevant cross-sectional regression in period t is

$$\tilde{y}_t - \hat{B}\hat{w}_t - \mathrm{vec}_{1:m}\{\hat{b}_i'\hat{b}_i\tau_i/2\} = \hat{B}\lambda_{t-1} + \eta_t, \qquad \text{(Appendix B.2)}$$

again estimated by GLS using $\widehat{var[\eta]}$ from above. As in this case the regressand is time-varying, our risk prices will be so, too. This corresponds to the approach of Gultekin and Rogalski (1985), who used returns and so did not have the convexity term. They got loadings from the classical factor analysis in the first step, and found that at least two factors were priced. They did not adjust for fitted factors, i.e., the term $\hat{B}\hat{w}_t$ was ignored, so their risk price estimates correspond to $\hat{w}_t + \lambda_{t-1}$ in our case. Thus, their estimates include the factor variation. In our case [Eq. (Appendix B.2)], the fitted factors reflect the portion of the centered data $\tilde{y}_t - \hat{\mu}$ related to the loadings $\hat{B}$, and λ_{t-1} picks up the part of $\hat{\mu}$ less the convexity term explained by $\hat{B}$. As $\hat{\mu}$ and $\hat{B}$ are time-invariant, λ_{t-1} should be so, too. The time series average of the fitted factors is zero by construction, and if the true but unobserved factor realizations had a nonzero average over the sample period, then this would be picked up by the estimated risk premiums. Thus, following Shanken (1992), the variance of the average factors constitutes a lower bound on the variance-covariance matrix of estimated risk prices. Estimated risk prices necessarily reflect the factor average, but in the Gultekin and Rogalski (1985) regression they actually reflect the full impact of the factors period by period, instead of the cross-sectional pricing. In Eq. (Appendix B.2), estimated risk prices explain cross-sectional pricing and could be entirely unrelated to the fitted covariance-generating factors, but the analysis also shows that our risk price estimates should be constant through time ($\hat{w}_t$ has exhausted the time variation). This is in contrast to our general state space approach, which allows estimating genuine time-varying risk prices and separate them from the covariance-generating factors.

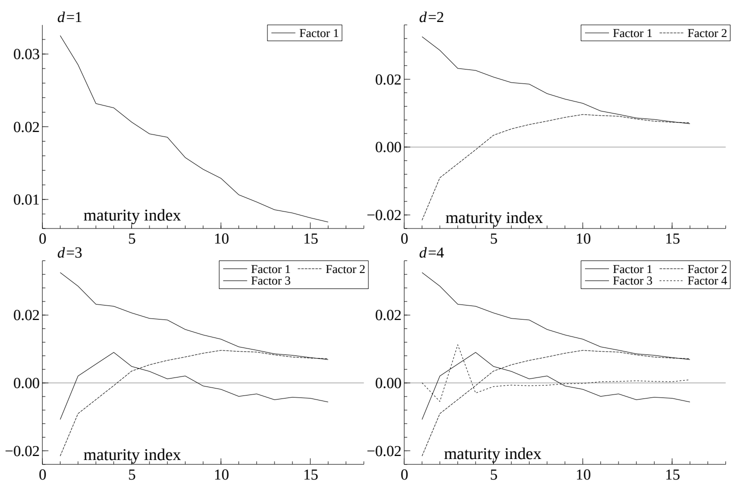
In Fig. B2 we show the time series of estimated risk prices λ_t from regression Eq. (Appendix B.2). Unlike the fitted factors, they are not centered around zero, and, as expected, the time variation is very little and mainly due to numerical issues (the interval on the vertical axis is very narrow, for each of the risk prices).

Table B1
Risk price estimates – Two-step PCA and GLS approach.

This table reports the estimated risk prices using a two-step approach. In the first step, a principal component analysis (PCA) is run on the demeaned slope-adjusted yield changes. In the second step, a cross-sectional regression is run to obtain estimates of the risk premiums. The table shows the estimated λ for $d = 1, 2, 3, 4$. To account for heteroskedasticity and cross-correlation, GLS is used. Asymptotic t-statistics are given below the estimates in parentheses, and * and ** denote significance at the 5% and 1% level, respectively.

Risk prices – Two-step PCA and GLS				
	Number of risk factors			
	$d = 1$	$d = 2$	$d = 3$	$d = 4$
Factor 1	−0.376	−0.363	−0.371	−0.373
	(−1.37)	(−1.16)	(−1.05)	(−1.06)
Factor 2		−0.107	−0.0904	−0.089
		(−0.344)	(−0.256)	(−0.253)
Factor 3			0.0634	0.0516
			(0.179)	(0.146)
Factor 4				−0.113
				(−0.319)

Panel A: Volatility functions (B) – PCA



Panel B: Covariance-generating factors (w_t) – PCA

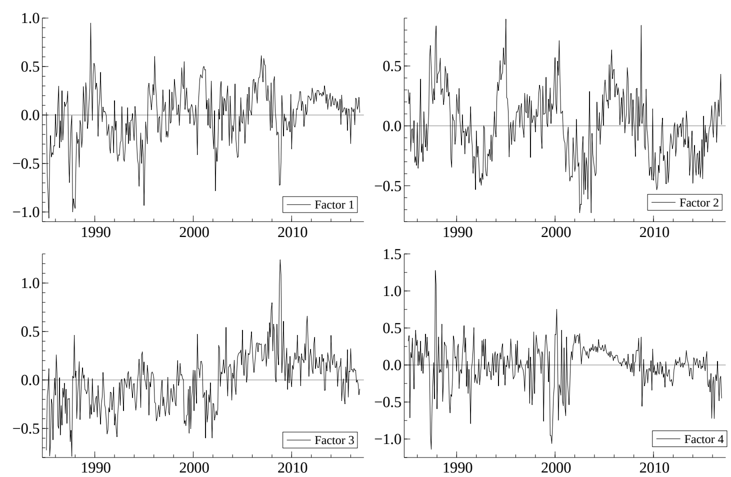


Fig. B1. This figure shows the volatility functions and estimated covariance-generating factors obtained by running a principal component analysis on the slope-adjusted yield changes. Panel A shows columns of the estimated B matrix by maturity for each number of factors $d = 1, 2, 3, 4$. The time series plots in Panel B show the estimated w_t from the $d = 4$ factor model.

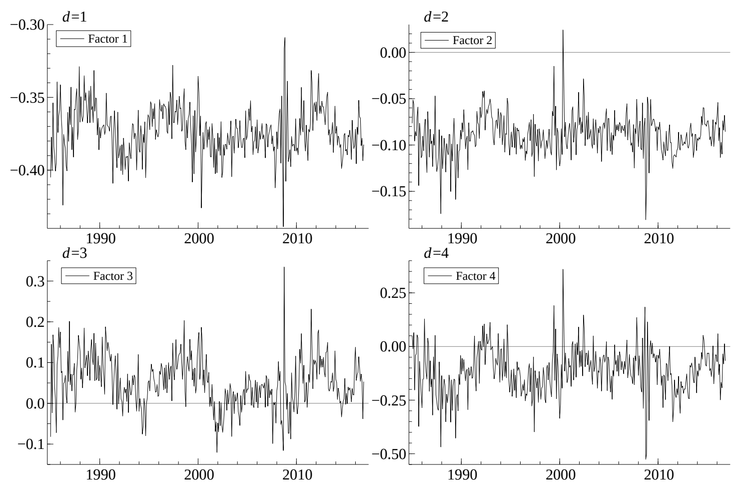


Fig. B2. This figure shows a time series plot of the estimated time-varying risk prices using a two-step approach. In the first step, a principal component analysis is run on the demeaned slope-adjusted yield changes. In the second step, a cross-sectional regression is run to obtain estimates of the risk prices. To account for heteroskedasticity and cross-correlation, GLS is used. The time series plots show the estimated λ_t for $d = 1, 2, 3, 4$.

Additional References (not in original paper)

Newey, W., West, K., 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Shanken, J., 1992. On the estimation of beta-pricing models. Review of Financial Studies 5, 1–33.