GRAPE Working Paper #113

Extracting Risk Free Interest Rate Expectations in a Less Liquid Government Bond Markets

Marcin Dec

FAME | GRAPE, 2026

Foundation of Admirers and Mavens of Economics
Group for Research in Applied Economics

Extracting Risk Free Interest Rate Expectations in Less Liquid Government Bond Markets

Marcin Dec
FAME|GRAPE

Abstract

This paper shows that in a less liquid government bond market, filtering term premia through a regression-based Adrian, Crump & Moench (ACM) framework yields risk neutral short rate expectations that match, and often rival, the accuracy of Survey of Professional Forecasters (SPF). Using monthly zero-coupon yields, we extract a model consistent risk free yield curve whose implied forward rates exhibit forecasting performance comparable to SPF paths across horizons up to three years. Crucially, these expectations can be generated daily, providing far higher frequency information than SPF's quarterly releases. We find that term premia are negligible at the short end but rise with maturity, and that the level factor—despite capturing most yield variance-does not command a price of risk. Cointegration tests indicate that SPF forecasts contain no incremental information beyond the filtered curve. The results highlight a practical advantage: once premia are removed, the yield curve becomes a reliable, high frequency source of monetary policy expectations suitable for policy analysis and market surveillance.

Keywords:
Term Premia Extraction, Risk Neutral Interest Rate Expectations, Yield Curve Decomposition, Survey of Professional Forecasters

JEL Classification:
E43, G12, G17

Corresponding author: :
Marcin Dec (m.dec@grape.org.pl).

Acknowledgments
The financial support of National Science Centre (grant UMO-2020/37/N/HS4/02202) is gratefully acknowledged.

Published by: FAME | GRAPE
ISSN: 2544-2473

1. Introduction and Methodology

The empirical motivation for this article is the observation established by Dec (2021), that the pure expectations hypothesis (PEH) tends to hold only locally in Poland, a market with a comparatively short modern time series and thinner secondary market liquidity than in the United States or the United Kingdom, and that its domain of validity may widen once one carefully filters the term-premium component that contaminates yield-curve-implied expectations, thereby disentangling the risk-neutral path of the short rate from compensation for bearing priced systematic risks. Accordingly, we formulate Hypothesis 1, namely that the PEH may obtain in a less-liquid market once term premia are correctly estimated and stress-tested, and, conditional on those estimates, we also ask to what extent term premia explain excess returns across the cross-section of maturities, while further formulating Hypothesis 2, which posits that professional forecasters' short-rate expectations do not contain incremental information beyond what is embedded in the term structure in such markets.

Our methodological approach is a deliberately pragmatic implementation of Adrian et al. (2012, 2013): rather than solving a full no-arbitrage likelihood with latent-state Kalman filtering, we rely on a three-stage regression algorithm that (i) extracts orthogonal state variables from a dense panel of zero-coupon yields via principal components analysis (PCA); (ii) estimates the mapping from states and one-step factor innovations to one-month excess returns across a fixed maturity grid; and (iii) recovers linear, state-dependent market prices of risk from cross-sectional return restrictions, which then permit the reconstruction of both a risky yield curve (with estimated prices of risk) and a risk-free curve (with prices of risk set to zero), the difference constituting the term-premium structure; we augment this pipeline with robustness diagnostics drawn from Malik and Meldrum (2014) and McCoy (2019), and we explicitly tailor choices-such as the short-rate proxy and the tenor grid-to the realities of a less-liquid data environment, thereby retaining the central ACM insights while keeping the computational burden light enough for extensive specification checks.

1.1 Data construction and the zero-coupon curve

We estimate continuously compounded zero-coupon yields using a preferred Nelson–Siegel–Svensson (NSS) specification on a monthly grid extending from the shortest available tenor to 10 years, using trading dates as close as possible to month-end to reduce microstructure noise and day-of-month seasonality; continuous compounding is chosen to maintain algebraic coherence with exponential-affine pricing and to keep state-space and return equations internally consistent, and, although compounding conventions do not affect the PCA itself, they greatly simplify the mapping between state dynamics, yield reconstructions and excess-return calculations in later steps. In our yield curve estimation, we follow a liquidity based weighting system proposed by Dec (2021). We assigned very high weight to the short end - anchoring the curve on the central bank bill's yield - while systematically excluding short-maturity bonds distorted by switch operations. To ensure that each instrument's importance reflected actual market conditions, we applied liquidity-based weights constructed as modified-duration multiplied by the sum of outstanding amount and turnover. We also kept the full maturity spectrum instead of truncating the long end, allowing the weighting system itself to down-weight thinner segments. Finally, we avoided equal-weight specifications, as they consistently generated the roughest and least reliable curve fits.

1.2 Dimensionality reduction: construction, interpretation, and macro-financial correlates

Let $y_t(\tau)$ denote the continuously compounded zero-coupon yield at calendar month t and maturity $\tau \in \mathcal{T}$, and collect the demeaned cross-section into the matrix $\widetilde{\mathbf{Y}}$. PCA solves the eigensystem $\hat{\Sigma}_Y v_k = \lambda_k v_k$ for the sample covariance $\hat{\Sigma}_Y = T^{-1}\widetilde{\mathbf{Y}}^\top\widetilde{\mathbf{Y}}$ and produces principal components $X_{t,k} = \widetilde{\mathbf{Y}}_t v_k$ which we standardize to unit variance, with the first K components retained for state-space construction. Variance decompositions show that PC1 explains roughly 96.2–97.7% of cross-maturity variation, PC2 explains 2.1–3.5%, and PC3 0.2–0.3%, with each higher component below 0.03%, a profile that confirms the well-known dominance of the level, slope and curvature trio while still leaving room for higher-order factors to matter for return pricing even if their variance shares are small. Factor loadings have the expected shapes, with the PC1 loading large, non-sign-changing yet not perfectly flat, PC2 crossing near four years-placing the pivot in the belly often regarded as less sensitive to immediate policy impulses-and PC3 partitioning maturities into short (≤1y), belly (1–6y), and long (>6y) segments, a tripartite structure plausibly linked to issuance patterns such as two- and five-year concentration, short-end switches, and thinner long-end supply.

Beyond geometric shapes, standardized factor time series exhibit economically meaningful co-movements with macro-financial variables as presented in Table 1. PC1 broadly mirrors CPI and PPI dynamics and is closely associated with the stance of monetary policy. It also correlates negatively with manufacturing output and retail sales, and positively with unemployment, the VIX proxy for global risk appetite, and the real effective exchange rate. PC2 is more cyclical and stationary, consistent with its common association with industrial production and labor-market slack and is negatively associated with stock-market growth.

Table 1. Correlation of factors from PCA with selected economic indicators

Indicator	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7
CPI	0.34			(0.20)	(0.21)	-	(0.21)
PPI	0.35	(0.22)	0.18	-	(0.24)		
Manufacturing (survey)		(0.19)		0.26			0.22
Confidence (survey)	0.49		(0.17)	0.24			0.21
Leading indicators		(0.29)		0.18	0.19		0.22
Industrial production		(0.34)		0.29			0.17
Manufacturing production	(0.84)	(0.17)			(0.25)		
Retail sales	(0.80)	(0.16)			(0.29)		
Unfilled vacancies	(0.65)	-	(0.18)	-	(0.36)		0.15
Unemployment rate	0.61		0.20	0.17	0.44		
Harmonised un. rate	0.59	0.15	0.19	0.28	0.42		
Harmonised un. rate (SA)	0.60	0.15	0.19	0.27	0.42		
USDPLN	(0.84)						
REER	0.71	0.26	(0.16)		(0.17)		
Stock market growth		(0.19)			0.26	0.18	
VIX (fear factor)	0.31	0.15		(0.31)	(0.19)		

Notes:

(1) only significant (95% confidence) coefficients are shown.

(2) source of time series: OECD, Main Economic Indicators – complete database

(3) data codes include: BVCICP02PLM460S, POLLOLITONOSTSAM, POLPRODMANMNISME, POLSPARTMSMEI, LMJVTTUVPLM647S, LMUNRTTPLM156S, LRHUTTTTPLM156N, LRHUTTTTPLM156S, CCUSMA02PLM618N, RBDPLEIS, SPASTT01PLM657N, VIXCLS.

(4) term premia are in columns; ACM model with K=5

Higher-order PCs, while harder to label, show distinctive if weaker correlational fingerprints-PC3 increases when the PLN depreciates, economic confidence deteriorates, and unemployment rises-altogether suggesting that PCA factors carry economically interpretable information even before one imposes any no-arbitrage structure or return-pricing restrictions.

1.3 The ACM framework in regression form: state dynamics, return equations, and prices of risk

We implement Adrian et al. (2012, 2013) approach to estimation of term premia through three OLS-based steps that together emulate the core exponential-affine logic while preserving tractability. In step 1, the standardized state vector $X_t \in \mathbb{R}^K$ follows a low-order monthly VAR, $X_{t+1} = c + \Phi X_t + \varepsilon_{t+1}$, with $\varepsilon_{t+1} \sim (0, \Sigma_\varepsilon)$ where c is typically very small due to standardization (or set to zero), and where we allow for short lags in Φ when slope forecasts are under consideration, given their stationarity. In the full $K = 7$ case the estimated transition matrix Φ appears as:

0.9964	-0.0069	-0.0111	0.0004	0.0107	-0.0020	0.0170
0.0119	0.9235	0.0152	0.0160	-0.0136	0.0269	0.0165
0.0502	-0.0075	0.8412	0.0123	0.0194	0.0478	0.0180
-0.0038	0.0303	0.0206	0.7457	-0.0185	0.1138	0.0318
0.0405	-0.0040	0.1476	-0.0503	0.5484	-0.1126	-0.0385
0.0044	0.0088	0.1092	0.1091	-0.1217	0.1273	-0.1323
-0.0068	-0.0308	0.0501	-0.0358	-0.0379	0.0365	0.0887

and the sub-matrices for $K \in \{3,4,5,6\}$ are close to the corresponding upper-left blocks of the $K = 7$ system, with only minor coefficient differences arising from re-estimation and lag reshaping-thus preserving persistence hierarchies across factors, with PC1 most persistent, PC2 stationary with moderate persistence, PC3 less persistent, and higher PCs approaching white noise.

In step 2, for maturities $n \in \{6,12,24,36,48,60,72,84,96,108,120\}$ months, we compute the one-month log excess return on the n-month zero-coupon bond as $rx_{t+1}^{(n)} = r_{t+1}^{(n)} - y_t^{(1)}$, with $r_{t+1}^{(n)}$defined by the log price change from $P_t^{(n)}$to $P_{t+1}^{(n-1)}$, and we do not smooth the short rate; instead we use the NSS-implied one-month zero coupon $y_t^{(1)}$, which ensures that the ultra-short tenor is measured consistently with the longer maturities and avoids auxiliary filtering that would be particularly problematic in a short sample. We then run the linear projection $rx_{t+1}^{(n)} = \alpha^{(n)} + \beta^{(n)\prime} X_t + \gamma^{(n)\prime} \varepsilon_{t+1} + u_{t+1}^{(n)}$ to identify exposures to predictable state components and contemporaneous factor shocks, stacking across maturities to obtain the loading matrix β that will be central for identification and the rank-deficiency tests that follow.

In step 3, we posit linear, state-dependent market prices of risk $\Lambda_t = A_0 + A_1 X_t$and exploit the cross-sectional restriction $\mathbb{E}_t[rx_{t+1}] \approx \beta \Lambda_t$, which under the log-linear stochastic discount factor implies that the vector of expected excess returns is spanned by factor innovations scaled by Λ_t. Estimating the rows of (A_0, A_1) by OLS on the stacked system and applying Wald tests on entire rows allows us to determine which factors are priced, and once (A_0, A_1) are in hand we can iterate the state dynamics to reconstruct a risky yield curve $y_t^{\mathrm{YC}}(\tau)$ and, by imposing $A_0 = A_1 = 0$, a risk-free curve $y_t^{\mathrm{RF}}(\tau)$, with the difference $TP_t(\tau) = y_t^{\mathrm{YC}}(\tau) - y_t^{\mathrm{RF}}(\tau)$ representing the term premium at horizon τ.

1.4 Identification, factor-count selection, and residual diagnostics

Identification rests on orthogonality of PCA factors, the inclusion of contemporaneous innovations in the return equations, and the linearity of prices of risk (compare: Table 2 below).

Table 2. Market prices of risk (Λ) for seven factor ACM model of the Polish interest rates

Factor / stat	λ0	λ1,1	λ1,2	λ1,3	λ1,4	λ1,5	λ1,6	λ1,7	WΛ	Wλ1
PC1	(0.016)	(0.006)	0.004	(0.005)	0.004	0.012	(0.001)	0.017	6.564	4.484
PC2	0.002	(0.032)	**(0.050)**	**0.064**	0.039	0.002	0.033	0.019	**15.689**	**15.680**
PC3	0.066	(0.041)	(0.011)	**(0.084)**	**0.101**	0.054	0.069	0.026	**19.825**	**17.110**
PC4	(0.002)	(0.018)	0.035	0.027	**(0.205)**	0.049	**0.144**	0.046	**30.584**	**30.582**
PC5	(0.006)	0.036	(0.029)	**0.128**	(0.042)	**(0.373)**	(0.003)	(0.002)	**44.069**	**44.060**
PC6	0.001	(0.030)	(0.021)	0.081	0.114	(0.020)	**(0.658)**	0.033	**86.231**	**86.231**
PC7	0.020	(0.011)	(0.001)	0.012	(0.064)	0.027	0.127	**(0.744)**	**97.870**	**97.799**
t-stat1	(1.442)	(0.540)	0.407	(0.485)	0.329	1.094	(0.109)	1.573	0.584	0.723
t-stat2	0.094	(1.221)	(1.917)	2.472	1.497	(0.077)	1.271	0.737	0.047	0.028
t-stat3	1.647	(1.018)	(0.275)	(2.106)	2.518	1.344	1.730	0.649	0.011	0.017
t-stat4	(0.047)	(0.372)	0.727	0.562	(4.307)	1.029	3.014	0.952	-	-
t-stat5	(0.095)	0.591	(0.492)	2.117	(0.698)	(6.205)	(0.045)	(0.035)	-	-
t-stat6	0.007	(0.413)	(0.293)	1.127	1.585	(0.274)	(9.051)	0.449	-	-
t-stat7	0.266	(0.149)	(0.676)	0.158	(0.850)	(0.358)	1.663	(9.679)	-	-

Notes:
(1) this table reports λ coefficient estimates paired with their t statistic value based on standard error computed as in ACM.
(2) Last two columns provide Wald statistics to test if the whole row of λ_k is statistically different from 0.
(3) Certain factor is not priced if Wald test shows we cannot reject this null hypothesis.
(4) Bold text is used to annotate significant coefficients (at 5%).
(5) λy,z for PC is a regression coefficient of factor in the price of factor risk.

We subject the stacked return loadings β' to Anderson's rank-deficiency test and apply Wald tests for unspanned or "useless" factors, consistently rejecting the null hypotheses at $p < 0.001$ for the Polish curve, which suggests that higher K improves in-sample fit even if the incremental variance shares are small, though we remain mindful that such gains can translate into deteriorated forecast performance; residual diagnostics confirm that yield-pricing errors are small and, for maturities beyond three years, approximately normal, and that return-pricing errors are also small in absolute magnitude, allaying concerns that the exponential-affine approximation fails in a less-liquid environment and supporting, instead, the view that a modestly enriched factor space captures the salient sources of priced risk.

2. Term-Premia Estimation

2.1 Priced risk in Poland: the unpriced level and sparse higher-factor structure

A central result in our cross-sectional pricing stage is that the level factor (PC1) is not priced in Polish government bonds for any $K \in \{3,4,5,6,7\}$ both the constant and all state-dependent coefficients associated with PC1 fail row-wise Wald tests, in sharp contrast to the U.S. evidence in Adrian et al. (2012), where the level portfolio is priced across specifications. By contrast, the slope factor and selected higher-order components are priced, with the only exception being curvature (PC3) in the $K = 3$ model, and, moreover, constant terms are almost never significant, save for a single intercept in the $K = 3$

specification, a pattern fully consistent with standardized inputs and with risk prices driven by state variation rather than fixed premia.

We interpret the absence of a level premium as a manifestation of Poland's shorter modern history and compressed yield range, which lack the deep structural episodes as for example: 1970s oil shocks, 1980s banking crises, the dot-com boom-bust, the pre- and post-2007 turmoil, that, in the U.S., have historically strengthened the correlation between elevated yield levels and heightened systemic risk. Absent repeated episodes that taught investors to associate high levels with high compensation, the level portfolio does not span priced risk in excess-return space, whereas slope and selected higher-order components, which capture curvature of macro-financial dynamics and issuance-related segments, do.

2.2 Coefficients and patterns across dimensions; implications for factor sufficiency

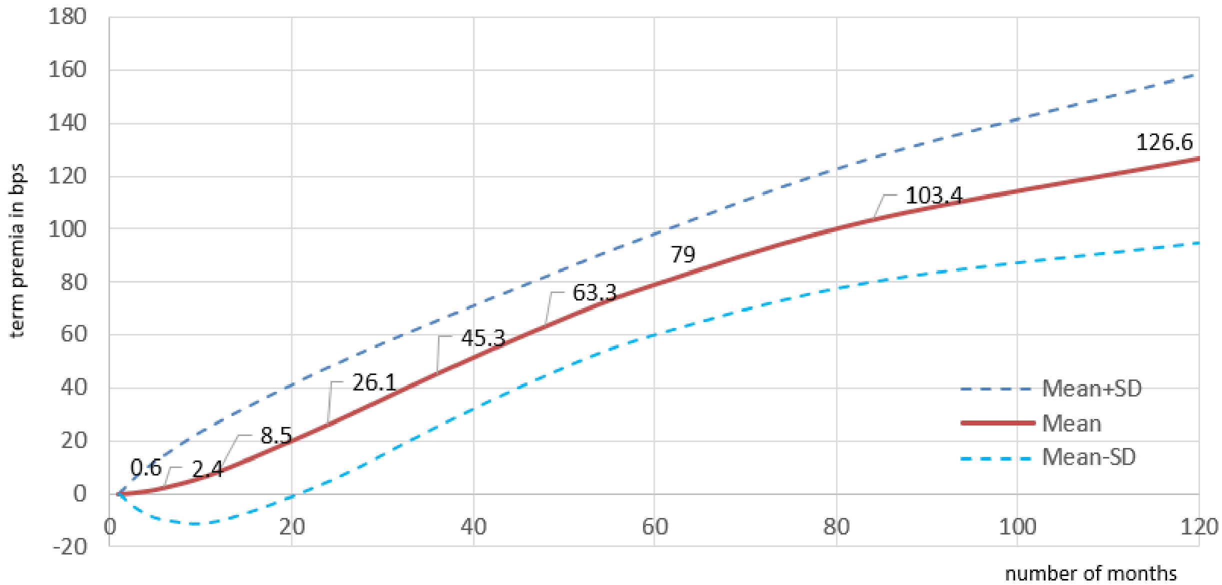
In the $K = 7$ specification, pricing is sparse and economically interpretable: PC2 risk is priced primarily through factors 2–3, PC3 through 3–4, PC4 through 4–5, PC5 through 3–5, while PC6 and PC7 are largely self-priced, and PC1 has no significant pricing components. For $K \in \{4,5,6\}$ the qualitative pattern and magnitudes persist, with only a minor loss of significance for the PC2-on-PC2 loading in $K = 5$, whereas $K = 3$ emerges as the weakest specification, with only slope risk priced-a finding that highlights the role of additional factors whose variance shares are small but whose exposures are informative about returns, particularly in the mid-curve where issuance practices and investor clientele effects may leave discernible footprints. Anderson rank and useless-factor tests, which we discuss below, are fully consistent with this qualitative picture.

2.3 Magnitudes and maturity profile of term premia

As presented in the Figure 1, averaging over more than 15 years of data (benchmark $K = 5$), term premia are single-digit basis points through one year, approximately 25 bps at two years, 80 bps at five years, and 125 bps at ten years.

Credible non-zero bands appear only beyond roughly the 2–3-year point, a configuration that dovetails with the pure expectations hypothesis tests which suggest that the very front end of the Polish curve can be treated as expectations-dominant, whereas the belly and the long end embed economically meaningful compensation for bearing factor risks. We emphasize that "risk-free" in our usage is a model-based notion-prices of risk set to zero-and not a statement about sovereign default probability.

Figure 1. The average term premia with one standard deviation bands by tenors for the Polish yield curve



2.4 Fit diagnostics, residual distributions, and factor-count decisions

Yield-pricing errors under ACM are smaller than the initial NSS fitting errors and, for maturities at or beyond three years, exhibit distributions close to normality; return-pricing errors are small in absolute terms, and the stacked loading matrix β' rejects rank deficiency at high significance, as do the Wald tests for useless factors, across all K considered, thereby arguing for the in-sample value of richer factor spaces. That said, the potential for overfitting in forecasting prompts us to treat $K = 5$ as a pragmatic benchmark that balances cross-sectional fit against out-of-sample stability, while retaining $K \in \{4,6,7\}$ as informative robustness variants and recognizing the clear inferiority of $K = 3$ for pricing and forecasting tasks.

3. Short-Interest-Rate Forecasts from the ACM Model

3.1 Forecast design, horizons, and loss functions

We conduct rolling pseudo-real-time forecast experiments for the one-month short rate, using two initial estimation windows-five and ten years-to calibrate the state dynamics and pricing stages, and we generate forecasts at horizons $h \in \{3,6,12,18,24,36,60\}$ months under two regimes: a risky regime (YC), which propagates the estimated prices of risk A_0, A_1through the state dynamics, and a risk-free regime (RF), which sets $A_0 = A_1 = 0$ while using the same state-space evolution, thereby eliminating the contribution of time-varying risk compensation to forecasted rates; naïve reference forecasts include a persistence (no-change) benchmark (hereafter: Per), and, in the level-forecast tasks for zero-coupon yields, an additional constant-short-rate assumption from the forecast origin; performance is assessed via RMSFE and MAFE and, importantly, we examine not only the means but also the dispersion of these losses across origins to evaluate forecast stability, and we split the analysis into level and slope components given the quasi-integration of the former and stationarity of the latter.

Table 3. Forecast errors of short interest rates in different models, 5-year window

Model variety	RMSFE							MAFE						
	3m	6m	12m	18m	24m	36m	60m	3m	6m	12m	18m	24m	36m	60m
Per	13.8	39.6	81.8	109.7	131.1	163.3	203.4	6.9	25.6	60.7	93.0	114.5	135.4	179.9
Imp, K=3	16.9	27.4	61.1	106.7	159.0	254.7	329.2	12.9	21.6	47.9	88.4	141.1	237.8	311.1
RF, K=3	21.1	40.6	81.3	120.6	156.7	215.2	273.9	16.0	33.0	66.0	99.2	133.2	200.7	269.6
Imp, K=4	16.6	28.3	65.1	106.8	154.4	246.8	351.4	12.8	21.5	52.8	89.3	136.9	229.2	334.0
RF, K=4	20.8	40.2	80.0	118.6	154.5	213.2	272.8	15.4	31.3	64.5	98.7	139.1	198.2	268.1
Imp, K=5	16.4	28.8	65.8	107.0	152.9	246.5	354.2	12.6	22.2	53.1	88.3	135.9	230.0	334.0
RF, K=5	19.2	37.5	78.3	118.3	154.9	214.1	273.0	14.1	28.7	62.5	97.5	132.1	198.6	268.7
Imp, K=6	16.5	28.9	65.8	107.1	153.1	246.4	354.8	12.7	22.2	53.0	88.9	135.6	229.9	334.0
RF, K=6	20.1	38.3	79.1	119.4	156.4	215.3	273.9	15.0	29.7	63.4	99.2	134.2	201.0	270.6
Imp, K=7	16.5	28.9	65.8	107.0	153.1	246.4	355.0	12.7	22.2	53.0	88.9	135.6	229.9	334.2
RF, K=7	21.3	39.4	79.8	120.9	159.1	218.5	276.1	16.0	31.9	65.4	101.0	137.1	205.6	273.5

Notes:
(1) various horizons are reflected in different columns h = {3, 6, 12, 18, 24, 36, 60}.
(2) Per means a benchmark of constant short interest rates, Imp stands for the forecasts implied by the risky yield curve and RF – by the ACM-derived risk-free curve.
(3) all errors in bps.

3.2 Headline patterns across horizons and specifications

Four robust patterns emerge from these exercises (see: Table 3):

(i) at short horizons up to 18 months-especially at 6 and 12 months-the risky YC forecasts dominate the persistence benchmark, whereas RF forecasts trail only slightly, an illustration being the $K = 4$ case at 12 months where the MAFE triplet (Per, YC, RF) is approximately 93, 89 and 98 bps, respectively,

(ii) at horizons of 24 months and beyond, RF forecasts become clearly preferable to YC forecasts, achieving lower average losses and, crucially, lower dispersion, which indicates that removing the stochastic term-premium component stabilizes medium-term projections,

(iii) increasing the factor count K only rarely degrades performance, and when it does, the deterioration is modest-typically 1–3 bps-and most evident in MAFE at 24–36 months,

(iv) the standard deviation of RMSFE and MAFE across forecast origins is markedly lower for RF at horizons at or beyond three years, signaling a tangible stability benefit from filtering out time-varying risk compensation.

3.3 Levels versus slopes; state generators and window length

For level forecasts of zero-coupon yields (up to 10 years) and horizons up to 24 months, the best ex post performance arises when the state generator is built from multiple low-order AR lags (1–3 months) and when the short rate is held constant from the forecast origin, a combination that accommodates the near-unit-root behavior of the level factor and resists the injection of transitory noise through fluctuating prices of risk. By contrast, slope forecasts benefit from VAR state generators with 3–6-month lags and from using the curve-implied short rate rather than freezing it, reflecting the stationarity and faster mean reversion of spreads and the value of cross-equation dynamics; across both tasks, ten-year estimation windows systematically dominate five-year windows, with especially visible gains for slope forecasts where additional history improves identification without unduly importing regime shifts, a difference that matters all the more in a short and relatively calm sample like Poland's.

3.4 Interpretation: horizon-dependent value of risk premia

The crossover in performance between YC and RF regimes is straightforward to rationalize: writing the risky forecast as the sum of the risk-neutral expectation and the expected future term-premium path clarifies that, at short horizons, time-varying prices of risk co-move with near-term cyclical indicators and therefore encode genuine predictive signal, whereas at medium horizons those prices of risk mean-revert, so their contribution mainly raises forecast variance without improving accuracy; setting $A_0 = A_1 = 0$ eliminates this mean-reverting noise, narrows the dispersion of projected paths, and thereby improves stability and average loss beyond 24 months, at the cost of occasionally discarding short-run information that the market briefly prices into the curve.

4. The Professional Forecasters' Challenge

4.1 Data reconstruction and alignment with model-implied paths

To confront model-implied short-rate paths with survey expectations, we use the National Bank of Poland's Survey of Professional Forecasters (SPF), available quarterly from 2011:Q3 onward, which elicits central projections for inflation and growth and, crucially for our purpose, includes additional questions on the NBP reference rate. Because the SPF reports average rates over non-aligned windows-selected future quarterly averages, calendar-year averages (for current and next years), and an average over the next five years-we first strip out already-realized information at the time of the survey (e.g., when 75% of the current calendar year is known at a 3Q survey, the submitted 2011 annual average must be de-biased for the known history), then map the residual forward-looking information into a sequence of true calendar-year averages starting at the survey quarter end, recognizing that the first "year" spans 0.25, 0.50, 0.75 or 1.00 years depending on whether the survey is conducted in 3Q, 2Q, 1Q or 4Q, respectively, and finally interpret the sequence as the one-year spot followed by four one-year forwards, which renders the SPF path commensurable with ACM- and YC-implied annualized short-rate trajectories. The out-of-sample window for accuracy comparisons covers 24 quarterly releases.

Formally, if the SPF provides an average rate $\bar{\imath}_{[t,t+1)}$ for the calendar year starting at the survey quarter end, but a fraction $\omega \in \{0.25,0.50,0.75\}$ of that year is already realized, we reconstruct the new information $\tilde{\imath}_{[t,t+1)}$ by solving $(1-\omega)\tilde{\imath}_{[t,t+1)} + \omega i_{\text{realized}} = \bar{\imath}_{[t,t+1)}$, and we apply analogous algebra to five-year averages to build a chain of year-ahead averages consistent with the survey's long-horizon constraints; we then aggregate ACM- or YC-implied monthly short-rate projections into comparable calendar-year averages and proceed to pairwise comparisons and cointegration analysis.

4.2 Descriptive features and horizon-by-horizon alignment

Descriptively, the SPF median path increases monotonically with the start date of the referenced year: the contemporaneous one-year average is around 2.15%, rising to 2.40% at one year ahead, 2.86% at two years, and 3.30% at three years, a shape that mirrors the YC- and RF-implied paths but is displaced upward by roughly 15 bps, as though forecasters treat the curve as embedding a maturity-dependent premium even after we filter term premia in RF, and as though judgmental adjustments wash out in the cross-section to a small positive bias at intermediate horizons. These features persist across choices of K and across alternative-yet reasonable-within-year weighting conventions in the reconstruction step.

4.3 Cointegration and incremental information tests

We test, horizon by horizon, whether the SPF median path contains information not already spanned by YC or RF. Johansen cointegration tests, conducted on the out-of-sample period and repeated for $K \in \{3,4,5,6,7\}$, show that SPF is cointegrated with YC for horizons up to two years and even more strongly with RF for horizons up to three years for $K \in \{3,4,5,6\}$, while for $K = 7$ the cointegration with RF extends to four years, indicating that SPF can be represented as linear combinations of curve-implied annualized expectations and does not supply orthogonal information beyond what is already embedded at higher frequency in the term structure (see: Table 4 below).

Table 4. Co-integration tests of three groups of time series: professional forecasters (PF), yield curve implied (YC) and ACM model implied risk free (RF)

x	Test result			p-value		
	x(PF) ~ YC	x(PF) ~ RF	x(RF) ~ YC	x(PF) ~ YC	x(PF) ~ RF	x(RF) ~ YC
1Yspot	1	1	0	0.044	0.005	0.081
2Yfwd	1	1	0	0.005	0.002	0.779
3Yfwd	0	1	0	0.873	0.028	0.916
4Yfwd	0	0	0	0.867	0.067	0.935

Notes:
(1) Engle-Granger tests for unit root in residuals from a regression x ~ Y, the null hypothesis Ho is that there in no co-integration.
(2) 1 - indicates rejection of Ho in favour of the alternative of co-integration, 0 - a failure to reject Ho
(3) RF and YC series are estimated using five factor ACM model and NSS curves with system of weights labelled 1.

While ex post error comparisons sometimes favor SPF within this short window, this apparent superiority is fragile because the data are quarterly, the sample is short, and much of it coincides with a secular decline in policy rates, all of which favor smooth, monotone survey paths.

5. Discussion

5.1 Why is the level factor unpriced in Poland?

That PC1 fails to price excess returns, despite dominating variance and mapping in the usual way to level, underscores the distinction between spanning co-movement and spanning compensation: with a shorter and calmer history, Poland offers fewer realizations in which high nominal yield states coincide with heightened macro-financial risk, thereby attenuating the prior that "high level = high risk compensation," so the market does not pay investors for holding the level portfolio; instead, priced risk resides in slope and selected higher-order components, which aligns with the empirical need for a hump-shaped risk-premium term structure that is negligible at the front end and rises into the belly and long end, exactly as the ACM extraction uncovers.

5.2 Forecasting regimes and the horizon-segmented strategy

The empirical crossover: risky dominance at 6–12 months, risk-free dominance at ≥24 months-provides a practical rule for users: in less-liquid markets, time-varying prices of risk carry exploitable signal at very short horizons but introduce variance at medium horizons; accordingly, tactical horizon users (e.g., dealers, short-term ALM) can rely on risky projections for ≤1 year, whereas strategic horizon users (e.g., policy analysts, long-horizon ALM and capital planning) should privilege risk-free ACM paths at ≥2 years, a division of labor that is robust to factor count and strengthened by using longer estimation windows and VAR-based slope dynamics.

5.3 Surveys as low-frequency mirrors of the curve

The cointegration of PF with both YC and RF-stronger with RF as K rises-suggests that survey medians reflect, perhaps implicitly, the information conveyed by the daily curve, possibly augmented by modest judgmental adjustments that wash out in the cross-section of respondents. Therefore, PF offers limited incremental content for tactical decision-making, and its role is best viewed as a low-frequency credibility check on the model-implied paths rather than as a driver of independent signal.

5.4 Limitations and extensions

The analysis is constrained by the finite sample (≈15.5 years) and the moderate amplitude of rate cycles in Poland. While robustness checks across K, lag structures, and window lengths show stable qualitative patterns, more data could refine inference on higher-order pricing and the stability of the risk-free forecasting advantage at long horizons; methodologically, likelihood-based extensions (e.g., Joslin et al.; Hamilton and Wu) would raise computational costs without obvious gain in this setting, whereas the regression-based ACM used here capitalizes on parsimony and transparency, which is desirable for surveillance and communication in markets where data are sparse and models must be explainable.

6. Summary and Conclusions

This paper demonstrates that once term premia are carefully filtered, the yield curve becomes a high-frequency expectations-extraction mechanism whose implied short-rate path is empirically indistinguishable from the SPF. In a less-liquid market such as Poland's, risk-neutral expectations derived from a regression-based ACM model achieve forecast accuracy comparable to quarterly expert forecasts, while being available daily and at virtually no additional informational cost. The SPF adds no incremental signal beyond what the filtered curve already embeds, underscoring that market prices-once stripped of risk compensation-encode the same macro-financial expectations that professionals reveal only intermittently.

These findings confirm the paper's main premise: term-premium filtering unlocks a practical, scalable method for extracting reliable monetary policy expectations at a far higher frequency than surveys allow. For policymakers, ALM practitioners, and market analysts, the risk-free yield curve thus offers a robust, timely, and transparent benchmark for expectation tracking in environments with limited liquidity and sparse macro data.

References

Adrian, T., Crump, R. K., & Moench, E. (2012). Efficient, Regression-Based Estimation of Dynamic Asset Pricing Models. SSRN Electronic Journal. doi: 10.2139/ssrn.2024297

Adrian, T., Crump, R. K., & Moench, E. (2013). Pricing the term structure with linear regressions. Journal of Financial Economics, 110 (April), 110–138. doi: 10.1016/j.jfineco .2013.04.009

Anderson, T. W. (1951). Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions. The Annals of Mathematical Statistics, 22 (3), 327–351. doi: 10.1214/aoms/1177729580

Annaert, Claes, Cleuster, Z. (2000). Estimating the Yield Curve Using the Nelson Siegel Model - A Ridge Regression Approach. Journal of Macroeconomics (2). doi: 10.2139/ ssrn.2054689

Dec (2021). Parsimonious yield curve modeling in less liquid markets. GRAPE Group for Research in Applied Economics, Working paper 52

Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. Journal of Econometrics, 130 (2), 337–364.

Hamilton, J. D., & Wu, J. C. (2012). Identification and estimation of Gaussian affine term structure models. Journal of Econometrics, 168 (2), 315–331. doi: 10.1016/j.jeconom.2012 .01.035

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24 (6), 417–441. doi: 10.1037/h0071325

Joslin, S., Singleton, K. J., & Zhu, H. (2011). A New Perspective on Gaussian Dynamic Term Structure Models. The Review of Financial Studies, 24 (3), 926–970. doi: 10.1093/ rfs/hhql28

Malik, S., & Meldrum, A. (2014). Evaluating the Robustness of UK Term Structure Decompositions Using Linear Regression Methods. SSRN Electronic Journal (518). doi: 10.2139/ssrn.2534455

McCoy, E. (2019). A calibration of the term premia to the euro area. (Vol. 8022) (No. September). doi: 10.2765/33831

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. The Journal of Business, 60 (4), 473. doi: 10.1086/296409

Pearson, K. (1901, Nov). LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2 (11), 559–572. doi: 10.1080/14786440109462720

Svensson, L. E. O. (1993). Term, inflation and foreign exchange risk premia. NBER Working papers.

Svensson, L. E. O. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994. NBER Working Paper (4871).