Bond Risk Premia and Discount Rate Variation in Brazil

Livio Maya*

Abstract

Long-duration public bonds denominated in local currency have been traded with liquidity in Brazil since the late 2000s. I use the thirteen years or so of now available data to study bond risk premia. As previously documented for the US, a single tent-shaped combination of forward rates predicts bonds' excess returns. R^2 surpasses 0.50. The same return-forecasting factor also predicts equity stock excess returns ($R^2 \approx 0.19$), and returns on the Real-Dollar carry trade ($R^2 \approx 0.17$), even after controlling for domestic interest, dividend-yield and US risk premium variation. The return factor is correlated with the MSCI index of emerging market risk premia and domestic interest. I input it along with term structure principal components on a exponential-affine model, and characterize model dynamics with a simple pair of co-movements involving the return factor and the level principal component. The latter I estimate to be correlated with domestic inflation, US bond risk premia and fiscal deficits, and whose shocks affect marginal utility in the model specification with the best fit.
Keywords: *Bond Risk Premia; Emerging Markets; Term Structure of Interest Rates; Exponential-Affine Models*
JEL Codes: *G12, G15*

1. Introduction

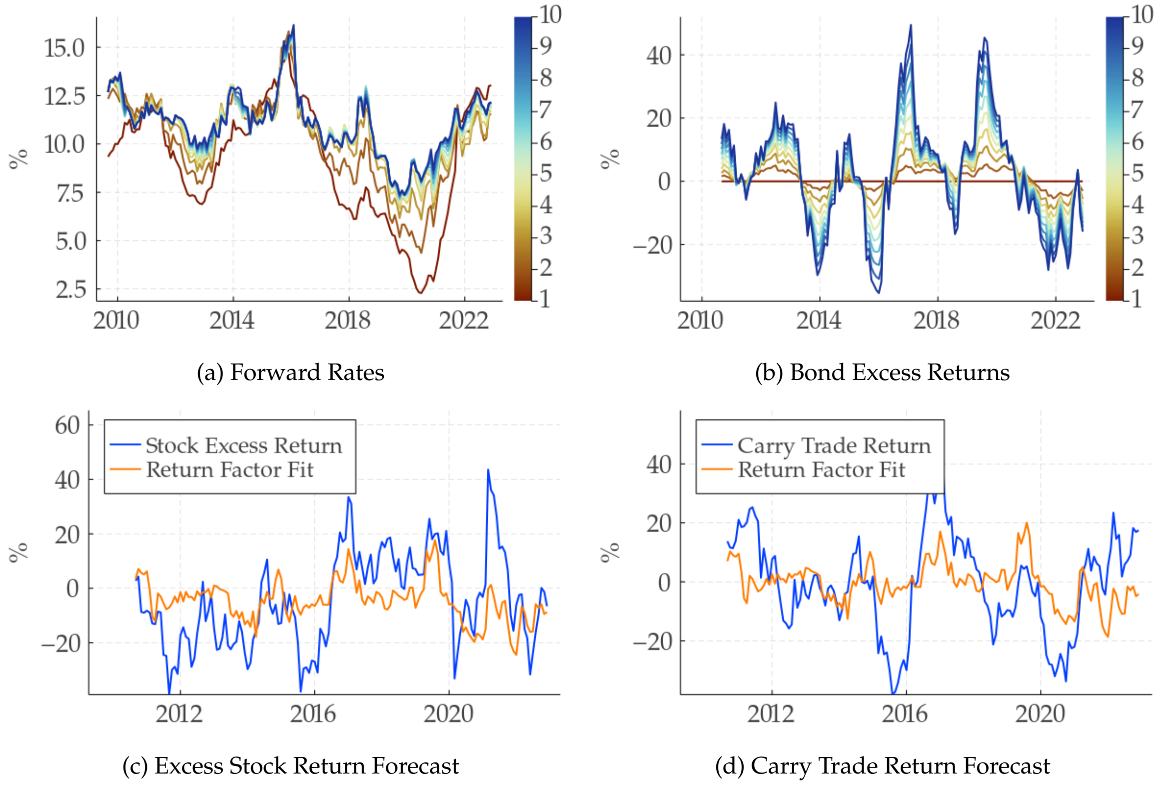
I study bond risk premia in Brazil. Using thirteen years of public bond price data, I run regressions of realized excess returns on forward rates, as Cochrane and Piazzesi (2005) do for the US case. Like them, I find that a single linear combination of yields, described by a tent-shaped profile of coefficients, captures the bulk of variation projected by the term structure in returns of bonds with maturities one through ten years. R^2s often surpass 0.40 (figure 1b). I report test results for the statistical significance of estimated coefficients using both asymptotic and small-sample estimators of their covariance matrices.

Like in Cochrane and Piazzesi's work, the linear combination of forward rates that predicts bond returns leads to a single "return-forecasting" factor. Besides the return of public bonds, I find that such factor also forecasts the return on Brazilian equity stocks ($R^2 \approx 0.19$, figure 1c) and returns on the Brazilian Real/US Dollar carry trade ($R^2 \approx 0.17$, figure 1d). Such predictability holds even after controlling for variation in Brazilian interest rates and dividend yields, and in US risk premium (proxied by the slope of its term structure, see Fama and Bliss (1987)). In all cases, p-values are lower or equal than 0.05.

To study the interpretation of the return factor, I set it as left-hand side variable on regressions against several domestic and international variables. The return factor is weakly correlated with the MSCI index and to interest rates, but disconnected from other domestic and international variables, such as US bond risk premium (as proxied by the slope of the term structure), domestic inflation and gross domestic product.

Also in accordance with results for the US, the return factor is not fully captured by traditional level, slope and curvature time series calculated from yields' (I use forward rates') principal components. While these three factors combined explain the average excess return with an R^2 of 0.30, the return factor alone reaches $R^2 = 0.47$. Moreover, in a regression with the four factors on the right-hand side, a χ^2 test fails to reject the null that the level, slope and curve factor coefficients are jointly zero. I conclude that the return factor summarizes the information contained in the term structure that is relevant to bond return predictability and hence risk premium measurement.

*Banco Safra. Email: liviomaya@gmail.com. I thank John Cochrane and Joaquim Levy for numerous useful comments.



(a) Forward Rates

(b) Bond Excess Returns

(c) Excess Stock Return Forecast

(d) Carry Trade Return Forecast

Figure 1: Data and predictability with a bond return factor

To study the dynamics of risk premia in Brazil, I then proceed to estimate an exponential-affine model of the term structure, as Duffie and Kan (1996) and Ang and Piazzesi (2003). I set as state variables the return factor along with term structure principal components (level, slope and curvature).

One can estimate factors' dynamics from time series moments, but, as Cochrane and Piazzesi (2009) point out, this approach leads to results that are highly sensitive to model specification.[1] Instead, I follow their procedure and estimate risk-neutral dynamics from the cross section of yields. Under a risk-neutral probability measure, forward rates equal expected future interest: $f_t^n = E_t i_{t+n-1}$. Hence, we directly observe a linear combination of state variables expected future path, which provides precise information about their dynamics. In symbols, if $i_t = \delta' x_t$ and $E_t x_{t+1} = \phi x_t$ (under the risk-neutral measure), then $f_t^n = \delta' \phi^{n-1} x_t$ and we can estimate ϕ without time series regressions. As the model can produce mean pricing errors as low as 0.06%, near the 0.05% produced by unrestricted OLS (the best linear estimator), I conclude it can precisely reproduce factors' dynamics under the risk-neutral probability measure.

I then recover the true probability measure using market prices of risk. Two assumptions reduce the number of prices of risk parameters, from twenty to two. I impose, first, that discount rates vary only due to variation in the return factor, an assumption justified by the findings of the first part of the paper. Second, I restrict the model's discount rate to price shocks to a single factor only. While I try different specifications, a priced *level* factor shock leads to the best model fit. The level factor is correlated with domestic inflation ($R^2 \approx 0.38$ in a linear regression), the US term spread ($R^2 \approx 0.35$) and primary surpluses ($R^2 \approx 0.23$).[2]

The estimated dynamics present a notable movement pattern. Variation in the level factor induces variation in the return factor (hence in risk premia) of equal sign. Higher level forecasts higher returns. Under the true probability measure, variation in the return factor then produces variation in the level

[1] Cochrane and Piazzesi exemplify the issue by showing how a usual VAR model and a co-integrated model of forward rates - two plausible statistical specifications - lead to wildly different forecasts of interest rates.

[2] Neumeyer and Perri (2005), Calvo et al. (2006) and Morelli et al. (2021) are examples of papers that relate international credit conditions to activity in emerging markets.

factor in the *opposite* direction. The system becomes stable: high level $\rightarrow$ high return $\rightarrow$ lower level. Under the risk-neutral measure, the second implication *fails*. The increase in risk premia does *not* bring down the level of yields, and so the system acquires a near-unitary root.[1]

This pattern, which is robust across multiple specifications, opens the door to interpreting events as the result of the co-movements of the state variables of the model. Positive shocks to inflation, surpluses or international risk premium, which typically lead to a higher level factor, forecast subsequent surges in Brazilian risk premia, which in turn predict a decline in the level factor. This is exactly the pattern observed in the 2015-16 recession. From the differences between true and risk neutral probabilities, we can also infer that the marginal utility of Brazilian bond investors is sensitive to combinations of structural shocks that provoke prolonged periods of elevated international risk premia or domestic interest, inflation and fiscal deficits. I make this point rigorous in the text.

The article offers two contributions to the existing literature. The first is to investigate if well-studied results about public bonds found in the US case, especially in terms of return predictability and the expectations hypothesis (Fama and Bliss (1987), Campbell and Shiller (1991), Bekaert and Hodrick (2001)), hold in a large, hopefully representative developing country. For example, I show that Fama-Bliss regressions fail to reject the expectations hypothesis, and how the model makes that consistent with predictability. The fact that my return-forecasting factor for nominal Brazilian bonds also predicts returns of equity stocks and currency portfolios extends the scope of the paper's contribution to include discount rate variation more broadly.

Studies of predictability in emerging economies are relatively scarce (Akgiray et al. (2016) and Devpura et al. (2021) are recent examples), as many markets lack the required data or liquidity. In the case of Brazil, following the end of hyperinflation in the mid-1990s, sequential administrations stretched the duration of public debt and changed its composition to encompass primarily local-currency bonds. We now have over ten years of market price data, provided by the Brazilian Financial and Capital Markets Association (ANBIMA), to study the questions above. To the best of my knowledge, this is the first research article to consider this dataset.[2]

The second contribution is to measure and characterize bond risk premia in Brazil, and risk premia more broadly given the high predictability I find in the stock and currency markets using the return factor. Knowledge of the expected path of risk premium at a given state of the economy is particularly informative about the marginal utility of the Brazilian investor. On a frictionless market, the first-order condition for the portfolio selection problem is $p_t = E_t(m_{t+1}x_{t+1})$, where p is the asset price, x represents its payoff, and the stochastic discount rate m is the investor's *marginal utility growth rate*. A risk-neutral probability measure π^* changes the actual measure π by "factoring in" the m: $\pi^* = \pi \times m$, state by state. By contrasting the economy's dynamics under π^* and π, we thus highlight the states in which marginal utility change the most, up to a shock orthogonal to payoffs (if $E_t(\epsilon x_{t+1}) = 0$, then $m + \epsilon$ is a discount factor too).

The difference between risk-neutral and actual probability therefore suggests that the marginal utility of Brazilian investors is sensible to prolonged periods of high inflation or international risk premia *that are not reversed by a rise of risk premia at home*. I interpret this finding to be consistent with macroeconomic literature arguing that the business cycle of emerging markets is affected by trend or long-lived shocks (as Aguiar and Gopinath (2006), Aguiar and Gopinath (2007) and Boz et al. (2011)), the effects of which cannot be reversed by a few years of risk premium variation.[3]

Finally, as a by-product of the exercise, I provide a decomposition of Brazilian forward rates into expected interest rates and term premia. Mid and long-term expected interest are volatile and explain a considerable share of the variation in forward rates. Expected interest does *not* quickly converge to

[1]As a consequence of the pattern I explain here, shocks to the level factor are very persistent under the risk-neutral measure. This is a property the model requires to generate the empirical fact that the first principal component explains most variation in forward rates.

[2]Papers studying the Brazilian term structure typically use swap data on the overnight lending rate (the DI rate). Datasets beginning prior to 2009 usually contain data on bonds with maturity up to two years. See, for instance, Tabak and Andrade (2003), Osmani and Tabak (2008), Almeida and Faria (2014), Almeida et al. (2015) and Faria and Almeida (2018).

[3]That the model fit improves when we set the level factor shock to be priced (be part of marginal utility) only increases the case for its economic significance.

the unconditional average. The term premium is countercyclical, but was considerably bigger in the 2015-16 crisis than during the early months of the COVID Pandemic. Also, term premia of bonds with different maturities do not move in lockstep. There is a term structure of risk premia.

2. Data

I use discrete-time language and, later, a discrete-time model. One period unit represents one year. Let p_t^n be the log-price in period t of a zero-coupon bond that pays after n years have passed. Let $f_t^n = p_t^{n-1} - p_t^n$ be the (log) forward rate at which one can lock today a one-period borrowing $n-1$ years ahead.[1] Let $i_t = f_t^1 = -p_t^1$ be the risk-free rate, which I also call the interest rate. Let $rx_{t+1}^n = p_{t+1}^{n-1} - p_t^n - i_t$ be the log of the *ex-post* return on the holding of a n-period bond from period t to $t+1$, in excess of the risk-free rate. I simply call them excess returns. Also, I define $\bar{rx}_{t+1} = (1/9)\sum_{n=2}^{10} rx_{t+1}^n$ as the average excess return.

The Brazilian Financial and Capital Markets Association (*ANBIMA*) collects market price data of nominal bonds issued by the Brazilian federal government. Like other countries, at any given point in time outstanding Brazilian bonds do not cover a full range of maturities, and not all bonds are zero-coupon. ANBIMA provides an estimate of the term structure of interest rates through a daily series of estimated load and decay parameters of a fitted Svenson (1994) model. The closest analogue to this estimation approach for the US is studied by Gürkaynak et al. (2007). ANBIMA re-calculates parameters every day to approximate the discount applied to the more liquid bonds outstanding in the Brazilian secondary market.[2] The Svenson (1994) model contains a total of six parameters. Therefore, the "data" follows from a highly non-linear six-factor model. From now on, I treat the fitted ANBIMA's curve as data, and highlight when this represents a caveat.

I use ANBIMA data to build monthly observations of the prices of bonds with maturities $n = 1, 2, \ldots, 10$ years. Price data then leads to a full term structure of forward rates, as well as realized excess returns. It covers the period Sep-2009 to Dec-2022, and each data point corresponds the last business day of the corresponding month. Note that returns are over a *one-year* holding of the bond. I do not consider lower holding periods.

Plots 1a and 1b show the data. Each curve represents a different maturity (in years), with the color bar on the right indicating which is which. The appendix provides the analysis of the unconditional moments estimated from the data.

2.1. *Fama-Bliss Regressions*

In the context of bond markets, return predictability manifests as a refusal of the *expectations hypothesis*, which one can define as the claim that forward rates reflect only the expected value of future interest. Indeed, the definitions above lead to the following relationship:

$$f_t^n = E_t i_{t+n-1} + \underbrace{E_t(rx_{t+1}^n - rx_{t+1}^{n-1}) + E_t(rx_{t+2}^{n-1} - rx_{t+2}^{n-2}) + \cdots + E_t(rx_{t+n-1}^2)}_{\text{Term Premium}_t^n} \quad (1)$$

Decomposition (1) holds *ex-post* (Cochrane and Piazzesi (2009) provide pretty art illustrating it), and it holds *ex-ante* in expected value. It says that a forward rate is given by the sum of the expected interest in the corresponding period and a *term premium*, which sums over time the differences in risk premia of bonds with consecutive maturities. If the expectations hypothesis holds, $f_t^n = E_t i_{t+n-1}$. Applying that to (1) iteratively leads to $E_t rx_{t+1}^n = 0$ for all n. Returns are unpredictable.

[1] I could alternatively work with yield-to-maturities which, in log, equal the average of the corresponding forward rates. I choose to work with the latter, as forward rates have a clearer connection to expected future interest. Using forward rates also leaves results more comparable to the closest literature for the US.

[2] ANBIMA provides market prices of different traded bonds, but not liquidity data, in that replication of the models' parameters is not feasible.

Maturity n	β_n	Small Sample std(β)	Asymptotic χ^2	p-value	R^2	Model β_n
Left-hand variable: Interest Growth						
2	1.64	0.39	8.97	0.00	0.34	1.10
3	1.55	0.54	5.15	0.02	0.28	1.27
4	1.52	0.62	7.09	0.01	0.26	1.43
5	2.19	0.64	34.16	0.00	0.54	1.48
Left-hand variable: Excess Return						
2	-0.64	0.39	1.37	0.24	0.07	-0.10
3	-0.35	0.50	0.21	0.65	0.01	0.12
4	-0.21	0.67	0.05	0.83	0.00	0.29
5	-0.20	0.87	0.03	0.86	0.00	0.44

Notes: small-sample standard deviations calculated from bootstrapping residuals from the simplified VAR $f_t^n = \sum_{j=1}^{12}(a_{n,j}^1 f_{t-j}^1 + a_{n,j}^{10} f_{t-j}^{10}) + \epsilon$. I average over 10,000 simulated paths with same length as the data (145 observations). The χ^2 statistic tests $\beta = 0$ using the asymptotic distribution. Spectral density matrix estimated using Newey-West correction with 18 lags.

Table 1: Fama-Bliss Regressions

Fama and Bliss (1987) regressions provide a straightforward test of the expectations hypothesis. Their empirical model is:

$$i_{t+n-1} - i_t = \alpha_n^i + \beta_n^i(f_t^n - i_t) + \varepsilon_{t+n-1}^n \tag{2}$$

$$rx_{t+1}^n = \alpha_n^x + \beta_n^x(f_t^n - i_t) + \varepsilon_{t+1}^x. \tag{3}$$

The expectations hypothesis implies $\beta_n^i = 1$ and $\beta_n^x = 0$. Past research has consistently estimated just about the opposite for the US market, $\beta_n^i = 0$ and $\beta_n^x = 1$, a fact attributed to slow adjustment of interest rates.[1]

Table 1 reports results with Brazilian data. I consider forward rates with maturities of two to five years. Since small sample issues are a big concern with only twelve years of data, I calculate standard deviations by bootstrapping residuals from a simplified VAR (see the notes on the table). I also report the χ^2 statistic for the test $\beta_n^i = 0$ or $\beta_n^x = 0$. I calculate it using a spectral density matrix estimated through the Newey-West approach, with 18 months of lags.

Results are the opposite of the usually reported for the US. They are closer to supporting, and even extrapolate in the direction of the expectations hypothesis. Estimates of β^i are greater than one in all cases. For $n = 2$ and $n = 3$, the difference is lower than one standard deviation. It is lower than two standard deviations in all cases. Results are economically significant: point estimates greater than one and the large R^2s suggest that forward spreads have predicted interest rate growth.

The bottom part of the table sets excess returns as the predicted variable. All estimated coefficients but that of $n = 2$ fall within one standard deviation away from zero. The χ^2 test fails to reject the null by a long margin, and R^2s are less than unimpressive.

In all, I conclude that Fama-Bliss regressions fail to reject the expectations hypothesis in the Brazilian market. A potential explanation lies in the high variability of the Brazilian interest rate (see the red line in figure 1a). The average absolute interest rate growth in the sample is 2.6%, a large value compared to forward rates, which typically range between 9% and 12%. In the Brazilian case, finding predictability demands that we go further than Fama-Bliss regressions. I later return to the issue and analyze results

[1] To see the connection, suppose an investor observes the forward spread $f_t^n - i_t$ and decides to purchase the n-year bond. In the following period, suppose that the interest rate stays the same, and that so does the term structure of forward rates, $p_{t+1}^n = p_t^n$. The investor's excess return then is

$$\begin{aligned} rx_{t+1}^n &= p_{t+1}^{n-1} - p_t^n - i_t \\ &= p_t^{n-1} - p_t^n - i_t \\ &= f_t^n - i_t \end{aligned}$$

and $\beta_n^x = 1$ becomes clear.

using the model.

3. Forecasting Returns

As Cochrane and Piazzesi (2005) do, I run regressions of bond excess annual returns on forward rates.

$$rx_{t+1}^n = \alpha^n + \beta_1^n i_t + \beta_2^n f_t^2 + \cdots + \beta_{10}^n f_t^{10} + \varepsilon_{t+1}^n \tag{4}$$

Forward rates are highly correlated, as they follow from a fitted factor model. Its non-linearity means we can still invert f's covariance matrix, but results scream the issue of multicollinearity.[1] When I run the regression with all or even a large subset of forward rates on the right-hand side of (4), estimated coefficients display a "W"-shape. I show the case with five forward rates ($n = 2, 4, 6, 8, 10$) in figure 2d. Each curve corresponds to the set of coefficients for a different maturity n, as indicated by the color bar.

The R^2s of regressions with all forward rates as explanatory variables vary between 0.50 and 0.57. I report them on the appendix, along with the ones for specifications with less forward rates. I do not adjust R^2s for model parameterization. It turns out we can stick to most predictability by keeping only three forward rates on the right-hand side. Forgoing additional terms substantially reduces R^2s , at least in some regressions.

I find that the combination of short, medium and long-term maturities yields the best overall fit. In my baseline specification, I use n =2, 6 and 10, which leads to R^2s in the range of 0.32 to 0.51. Define $F_t = (f_t^2, f_t^6, f_t^{10})'$. The unrestricted model therefore is

$$rx_{t+1} = \alpha + \beta F_t + \varepsilon_{t+1} \tag{5}$$

where rx stacks the returns of bond with maturities two through ten, α (9×1) and β (9×3) are the accordingly stacked coefficients.

Figure 2a display the estimated β. The "tent" shape of the coefficients and the fact that they appear to be scaled versions of each other reminds us of the analogous results found by Cochrane and Piazzesi (2005). As they point out, this profile of estimates suggests that one can keep most return forecasting power using a single factor rather than three. Such factor then captures most time-variation in the expected returns of bonds of *all* maturities.[2] There is nothing special about my choice of right-hand side forward rates: figure 2c shows that similar choices lead to the same "tent"-shaped profiles of coefficients.

Algebraically, we need to find a 9-entry vector q and a 3-entry vector γ such that $\beta \approx q\gamma'$. Then, the unrestricted model (5) becomes a one-factor, restricted model:

$$rx_{t+1} = \alpha + q\left(\gamma' F_t\right) + \varepsilon_{t+1}, \tag{6}$$

in which vector q stores the linear regression coefficient of each excess return on the single factor $\gamma' F_t$.

There are many reasonable ways to compute q and γ and address the fact that they are only defined up to scale. I follow the procedure Cochrane and Piazzesi use in their first article and normalize the *average* coefficient q_n to be one: $(1/9)\sum_{n=2}^{10} q_n = 1$. This is convenient because it allows me to, first, compute the γ's by regressing $\bar{rx}_{t+1}$ on F_t:

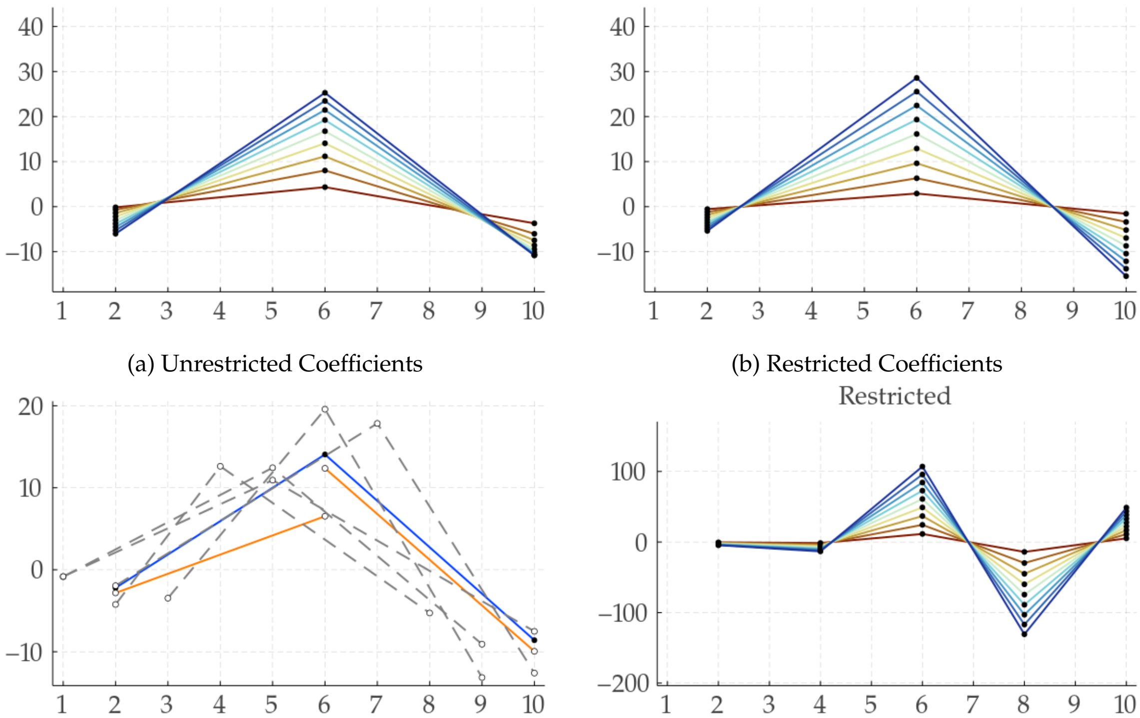
$$\bar{rx}_{t+1} = \bar{\alpha} + \gamma' F_t + \bar{\varepsilon}_{t+1} \tag{7}$$

and, then, estimate q by running OLS on (6) given γ. Figure 2b plots the restricted coefficients, that is, $q\gamma'$. It is visible that the restricted model produces a profile of coefficients that is similar to the unrestricted model.

Table 2 reports additional results. Panel A focuses on the first-stage regression (7). I divide it into two sections. The first section reports estimates and standard deviations in the baseline case. The

[1] Cochrane and Piazzesi (2009) document the same issue for the US, using fitted-model data from Gürkaynak et al. (2007). Their 2005 paper uses the Fama Bliss dataset (from CRSP), an analogous to which is sadly not yet available to Brazil.

[2] I assume the time-t information set to be the bonds' price data up to time t. The unrestricted model (5) and its restricted version (6) implicitly assume a Markovian structure, that is, that information prior to time t does not improve return forecasts.



(a) Unrestricted Coefficients

(b) Restricted Coefficients

(c) Unrestricted (rx^5, various F)

(d) Using five yields: multicollinearity

Notes:Panels 2a and 2d display the estimated coefficients of unrestricted regressions $rx_{t+1}^n = \alpha + \beta' f_t + \varepsilon_{t+1}^n$. Each curve corresponds to the coefficients for the regression on a different n, as indicated by the color bars. Panel 2c looks only at the regression of the five-year bond return rx^5 and plots the unrestricted coefficients for various choices of right-hand forward rates. The blue cuve is the baseline choice $F = (f^2, f^6, f^{10})$; the red curves use the two pairs (f^2, f^6) and (f^6, f^{10}). Panel 2b shows the coefficients of the regression with the restriction $\beta = q'\gamma$ (see text).

Figure 2: Return-Forecasting Regression: Coefficient Estimates

Panel A. First Stage Regression: $\bar{rx}_{t+1} = \bar{\alpha} + \gamma' F_t + \bar{\varepsilon}_{t+1}$						
	γ_2	γ_6	γ_{10}	$\chi^2(3)$	R^2	R^2 (t<2020)
Time-t forward rates						
Estimate	-3.00	15.97	-8.63		0.47	0.59
Std (Asymptotic)	(1.84)	(2.78)	(3.54)	41.12		
Std (Small Sample)	(1.43)	(4.00)	(4.00)	36.16		
One-month lagged forward rates						
Estimate	-2.38	15.40	-8.93		0.44	0.54
Std (Asymptotic)	(1.96)	(2.98)	(3.01)	45.13		
Std (Small Sample)	(1.36)	(4.02)	(4.03)	30.85		

Panel B. Second Stage Regression: $rx^n_{t+1} = \alpha_n + q_n(\gamma' F_t) + \varepsilon^n_{t+1}$						
n	q_n	Asymp. std	SS std	R^2	Unrest. R^2	χ^2 (3), $\beta = q\gamma'$
2	0.18	0.05	0.05	0.24	0.32	4.91
3	0.39	0.06	0.06	0.33	0.36	2.51
4	0.60	0.04	0.05	0.39	0.40	1.05
5	0.81	0.03	0.03	0.43	0.44	0.36
6	1.01	0.01	0.01	0.46	0.46	0.06
7	1.21	0.02	0.02	0.48	0.48	0.01
8	1.41	0.03	0.04	0.49	0.49	0.12
9	1.60	0.05	0.06	0.49	0.50	0.37
10	1.79	0.08	0.08	0.50	0.51	0.74

Notes: results from the restricted model (6). "Small Sample" covariance matrix calculated from 10,000 simulation paths computed from simplified VAR with bootstrapped residuals. See notes to table 1. "Asymptotic" covariance matrix calculated from GMM formulas for the two-step estimation procedure. R^2s are *not* adjusted for model parameterization. R^2 ($t < 2020$) indicates R^2 for the subsample ending in December 2019. The χ^2 column shows the test statistic for the hypothesis $\gamma = 0$.

Table 2: Return-Forecasting Regressions

estimates of γ are consistent with the tent shape. The black line in figure 1b plots the fitted value of the average return rate and illustrates the regression's large R^2 = 0.47.

One could worry that predictability follows from pronounced movements in the term structure during the COVID pandemic era, which represents a large shock in a thirteen-year database. This is not the case. Table 2 reports the R^2 for the regression run after excluding data from 2020 onward. R^2 goes up, not down.

I compute the covariance matrix of parameter estimates using two methods. In the first one ("asymptotic"), I estimate the spectral density matrix using the Newey-West formula with 18 lags and calculate the GMM covariance matrix.[1] The second method ("Small Sample") uses the same procedure explained in the Fama-Bliss regression (see notes to table 1). I compute the sample covariance matrix of estimates from ten thousand simulations of term structure prices, generated by a simplified VAR with bootstrapped residuals.

The table reports the χ^2 statistic for the hypothesis $\gamma = 0$ (no joint significance), which represents the hypothesis of time-invariant risk premia.[2] The one-percent critical value is 11.34. Using either method to calculate estimates' variance, the test statistic surpasses 30. We can confidently reject the hypothesis of time-invariant risk premium and the expectations hypothesis.

Panel B focuses on the second-stage regression. I report in the table the R^2s from the unrestricted model (5) for comparison. They vary from 0.24 to 0.48. The difference in forecasting power among bonds with different maturities is far more pronounced in the Brazilian case then in the US, as Cochrane and Piazzesi (2005) document. In addition, χ^2 tests of the joint significance of the β_n parameters (equation by equation, not reported) reject the null $\beta_n = 0$ in all regressions. The lowest value I find for the test statistic is 32.5. The table does report the χ^2 test for the restriction $\beta_n = q_n\gamma$. Again, the statistic only

[1] In this case, GMM estimates of the covariance matrix of q differ from OLS estimates, as parameter vector γ enters the moment condition $E[\varepsilon^n_{t+1} \times (\gamma' F_t)] = 0$ that pins down q. I approximate the derivatives of moment conditions using forward differentiation.

[2] Accepting the expectations hypothesis test here requires the inclusion of the constant term, which I do not. On the other hand, *rejecting* the time-invariant risk premium does imply rejecting the expectations hypothesis.

tests the joint significance of the equality *in a given equation n*. The test is *not* analogous to Cochrane and Piazzesi (2005), table 6, where the authors test for the joint hypothesis across all parameters *of all equations*. Testing the restriction equation by equation leads to a failure to reject the null $\beta_n = q_n \gamma$ for all n (the 10% critical value of the χ^2 distribution with three degrees of freedom is 6.25).

Focusing on the second-stage restricted regression itself, loadings of each bonds' return on the single factor q_n increase almost linearly, like in the US case. All estimates are highly significant individually, regardless of how I calculate standard deviations. The table shows that most variation in excess returns explained in a model with three factors is captured by the single return-forecasting factor $\gamma' F_t$. R^2s vary from 0.24 to 0.50.

Measurement Error. I follow Cochrane and Piazzesi (2005) and test for measurement error bias by running the return-predicting regression (7) with lagged forward rates on the right-hand side. Return-forecasting regressions that use price data on the right-hand side are always subject to having results biased by measurement error. If the price is inaccurately measured to be too high today and the error vanishes tomorrow, we observe a higher *ex-post* return just due to this error. If, however, measurement error is uncorrelated over time, the same return-forecasting regression that uses *lagged* forward rates on the right-hand side should be immune to measurement error bias. We should observe similar estimates.

This is precisely what the second section of Panel A shows. I lag forward rates a month and run regression (7). The model becomes $E_t \bar{rx}_{t+1} = \bar{\alpha} + \gamma' F_{t-1/12}$. The table reports the same tent-shaped profile of γ, with standard deviation and hypothesis testing conclusions similar to the baseline case, and only slightly lower R^2 . I try the same exercise with lags of two and three months - not reported - and find similar results. I also do not replicate Cochrane and Piazzesi's simulations of an expectations-hypothesis model with measurement errors (figure 4 of their paper). They show that such case would lead to estimates that form "step" functions, not the tent shapes that both studies find. In conclusion, we cannot attribute the main result of return predictability to measurement error.

Scaling. Define the return-forecasting factor to be $z_t = c_0 + c_1 \gamma' F_t$. I choose c_0 so that z_t has a sample average of zero and c_1 so that its sample standard deviation is the same as the interest rate f^1 (about 0.031). Mean/variance normalization of the return factor (which I call z from now on) is useful to interpret the coefficients I estimate next. It does not change its ability to forecast bond returns.

3.1. *Equity Stock Returns and Currency Carry Trade*

If the return factor successfully captures time variation in the risk premia of Brazilian public bonds, the next thing to ask is whether it also captures variation in the premium of other portfolios. It does. In this section, I regress the excess return of Brazilian equity stocks and returns to the Brazilian Real/ US Dollar carry trade on z_t.

My proxy for Brazilian stock returns r_t^S is the annual (log) price variation of the Ibovespa index.[1] The excess stock return is $rx_t^S = r_t^S - i_t$. Let e_t be the log of the price of one US Dollar in units of Brazilian Reals. The return on the currency carry trade is the difference between Brazilian interest rate (i_t) and US interest rates (i_t^*), minus the depreciation of the Brazilian Real with respect to the US dollar: $rx_t^E = i_t - i_t^* - \Delta e_{t+1}$ (the carry trade portfolio is a zero-price portfolio, hence the "excess return" notation rx).

For the US interest rate, I use the one-year Treasury Bill Secondary Market Rate, downloaded from the St. Louis Fed's FRED database. Ibovespa and exchange rate data I collect from the Brazilian Institute of Geography and Statistics and the dividend yield series from Bloomberg. Figures 1c and 1d plot rx^S and rx^E.

Panel A of table 3 reports the results. All regressions use an intercept term. The numbers in parenthesis are the p-values of the tests of individual significance (below coefficient estimates) or joint significance (under the "χ^2 Tests" columns). In all cases, I calculate the covariance matrix of coefficients' estimates using the Newey-West correction method, with 18 lags.

[1] Ibovespa is the main index of Brazilian stock prices. By construction, its growth rate includes stock price variation, dividend payments and stock buybacks.

Panel A. Stock and Carry Trade Returns

	Right-Hand Side Variables				χ^2 Tests			R^2
	z	i_t	dp_t	ts^{US}	All	Not z (3)	PCs (3)	
Stock Excess Return	2.34				15.3			0.19
	(0.00)				(0.00)			
	3.31	-0.99	-1.47	-9.91	47.2	22.4	10.8	0.41
	(0.00)	(0.53)	(0.61)	(0.04)	(0.00)	(0.00)	(0.01)	
Stock Return	2.75				22.7		5.4	0.27
	(0.00)				(0.00)		(0.14)	
Carry Trade Return	2.15				4.8			0.17
	(0.03)				(0.03)			
	1.74	0.71	-0.44	-2.37	8.6	1.1	9.02	0.20
	(0.05)	(0.67)	(0.88)	(0.78)	(0.07)	(0.78)	(0.03)	
Currency Depreciation	-1.77				3.8			0.13
	(0.05)				(0.05)			
	-1.82	0.26	0.75	-1.52	7.8	0.21	7.72	0.13
	(0.03)	(0.87)	(0.79)	(0.85)	(0.10)	(0.98)	(0.05)	

Panel B. Bond Return Predictability.

	Right-Hand Side Variables				χ^2 Tests	R^2
	z	Level	Slope	Curve	Not z (3)	
Avg Excess Bond Return	2.29					0.47
	(0.00)					
		0.81	1.71	2.23	13.1	0.30
		(0.01)	(0.16)	(0.18)	(0.00)	
	2.42	-0.01	-0.50	0.70	0.99	0.47
	(0.00)	(0.97)	(0.67)	(0.66)	(0.80)	

Notes: p-values indicated in parenthesis. Left-hand side variable indicated on the first column. Estimates covariance matrices calculated using Newey-West correction method with 18 lags. The "All" column shows the χ^2 statistic for the null hypothesis that all (slope) coefficients are equal to zero. The "Not z" column tests if the coefficients of all r.h.s variables but z are zero. The "PC" column in Panel A changes the set of r.h.s variable to be the return factor z and the principal components (level, slope, curvature; same variables used by Panel B) and tests if the estimates of the principal components' coefficients are jointly equal to zero.

Table 3: Return Factor z Forecasting Power

Regressions of the stock excess return and of the carry trade return share most results. The return factor z forecasts both. Coefficients are highly statistically significant, with p-values equal to or lower than 5%. They are also *economically* significant. My estimates indicate that, for a change of 1% in z_t, which varies as much as the economy's interest rate, investors expect a 2.3% increase in the stock excess return and a 2.15% increase in the carry trade return.

The model explains a lot of the variation in realized returns as well. R^2s are respectively 0.18 and 0.28. Cochrane and Piazzesi (2005) report an R^2 of 0.07 for stock return regression in the US case. Figures 1c and 1d plot the fitted values.

The table also reports results for the gross equity stock return r^S and currency depreciation Δe, to check if predictability follows from variation in the risk-free rates rather than variation in the risky components of each return. Again, the return factor maintains its forecasting ability, with $R^2 = 0.27$ and 0.13. Again, statistically and economically significant estimates.

Estimated coefficients appear to have the right sign as well. One can understand an equity share as a perpetuity with dividend risk. The stock return coefficient 2.34 is only slightly greater than the 2.29 we find for the average bond return, but becomes 3.31 after controlling for other factors. We can interpret the difference as additional premium for holding a "bond" with stochastic cash flows. As for the carry trade return, its only source of risk is currency variation.[1] The positive sign of the coefficient indicates that, at the same time investors demand a higher return from Brazilian stocks and bonds, they also demand a higher return from holding Brazilian currency ($\Delta e < 0$ means the Real appreciates compared to the US Dollar). This is exactly what one expects from a developing country whose currency does not have the "flight-to-quality" quality (as in Cho et al. (2016)).

Finally, I also run the same regressions using additional controls on the right-hand side. I include the Brazilian interest i_t, the Ibovespa dividend-yield dp_t and the US term spread ts^{US}. The latter I define to be the difference between the yields of the five and one-year T-bills. Dividend yields are known to forecast stock returns in the US[2], and the term spread I include as a proxy for international risk premia.[3]

In all cases, the return factor coefficient remains statistically and economically significant. In the case of the carry trade and currency depreciation, the coefficients multiplying the three controls are statistically equal to zero, both individually and jointly (the "Not z" column reports the χ^2 test statistic for null that these coefficients are jointly zero). The return factor drives out the effect of the other variables.

In the case of the stock excess return regression, the coefficient multiplying the US term spread is statistically significant at a 5%-sized test. The loading on the return factor is even higher compared to the univariate model, while that of the other two predicting variables are statistically zero. We gain a lot of predictability too: R^2 jumps to 0.41. A natural conjecture is that the US term spread captures variation in the market price of a risk factor specific to equity cash flows, given that it does not show up on the carry trade return equation.

3.2. *Principal Component Factors*

Cochrane and Piazzesi (2005) show that the information contained in their return factor cannot be summarized by the information contained in the traditional level, slope and curve factors. The same is true in the Brazilian case. While the first three principal components capture 99.5% of term structure variation, they do not capture the same share of the return factor's variation.

Panel B of table 3 shows that. I again regress the average excess return $\bar{rx}_{t+1}$ on z, but I now include level, slope and curvature factors on the right-hand side:

$$\bar{rx}_{t+1} = \alpha + \beta z_t + \psi_1 \text{ level}_t + \psi_2 \text{ slope}_t + \psi_3 \text{ curvature}_t + \varepsilon_{t+1} \tag{8}$$

[1] I do not discuss credit risk. Both the Brazilian and the American governments have most of their sovereign debt issued in local currency, in that "defaults" take place via inflation and, thus, changes in the relative value of each currency. These changes are by definition reflected by the nominal exchange rate.

[2] See Fama and French (1988), Campbell and Shiller (1988), Cochrane (1992), and many other since.

[3] Fama and French (1989) conclude that the term spread forecasts both stock and bond returns in the US.

The first row of the panel considers only z as explanatory variable. R^2 = 0.47 matches that of the first-stage regression in table 2, as it should. The second row considers only the principal components. Combined, they reach R^2 = 0.30. I do not adjust R^2s for model parameterization - they reflect only summed squared errors. On its own, the return factor explains the average return on Brazilian nominal bonds better than the traditional factors combined.

Another key point follows from the last row, which considers the four explanatory variables combined. In that case, coefficient estimates of the principal component factors are individually and jointly statistically equal to zero. The p-value for the χ^2 test is 0.80, an overwhelming failure to reject the null $\psi = 0$. I conclude that, given the return factor z, the three traditional term structure factors offer no additional forecasting power to bond returns.

I also test whether these factors improve predictability of the other portfolios considered in Panel A. To do this, I run a regression similar to (8), but replace the dependent variable. Column "PCs (3)" reports the same χ^2 statistic testing $\psi = 0$ along with its p-value. We cannot conclude that the return forecasting factor captures the whole predictability embedded in bond prices in the case of these assets. Unreported point estimates suggest that the level factor adds relevant information in forecasting both stock excess returns and returns on the carry trade.

4. Risk Premia Dynamics

The main point of the last two sections is that the return-forecasting factor z_t contains relevant information regarding the risk premium of different assets in the Brazilian economy. I now estimate a parametric model to characterize its conditional distribution.

4.1. The Exponential-Affine Model

Following Ang and Piazzesi (2003) and Ang et al. (2007), I consider the homoscedastic, discrete-time version of the multifactor exponential-affine model presented by Duffie and Kan (1996). Because the model has already been extensively studied by the existing literature, I spend little time describing the equations. First-time readers are encouraged to check the above references or the appendix, where I provide a full derivation of the model.

I use four factors. The first factor is the return-forecasting factor z. The next three are the first three principal components of the term structure (level, slope and curvature). I normalize all factor to have zero sample mean:

$$X_t = \begin{bmatrix} z_t \\ \text{level}_t \\ \text{slope}_t \\ \text{curve}_t \end{bmatrix} \quad \text{(demeaned)}$$

You can find their time series in the appendix. Factors evolve according to

$$X_t = \Phi X_{t-1} + e_t, \tag{9}$$

where $e_t \sim \mathcal{N}(0, \Sigma)$ is independent over time and captures the shocks hitting the economy. Model (9) implies that the populational mean of X is zero. Implicitly, I assume that the sample averages of the four factors coincide with their populational means. I later relax this assumption and find better fit, but results that lead to similar conclusions.

I concluded in the previous section that, given the return-forecasting factor z, the three principal components offer no additional return predictability. Hence, we would like the model to replicate the empirical regression

$$rx^n_{t+1} = \alpha_n + \beta_n z_t + \epsilon_{n,t+1}. \tag{10}$$

The exponential-affine model allows us to reverse-engineer a stochastic discount factor $\exp\{m_{t+1}\}$

that achieves that. In particular, assume

$$m_{t+1} = -\delta_0 - \delta_1' X_t - (1/2)\lambda_t' \Sigma \lambda_t - \lambda_t' e_{t+1}$$
$$\lambda_t = \lambda_0 + \lambda_1 X_t. \tag{11}$$

The following proposition summarizes the properties of the models' solution that I use later.

Proposition 1 (Solution to the exponential-affine model)**.** *The solution to the exponential-affine model involves log prices and forward rates that are linear on the factors:*

$$p_t^n = A_n + B_n' X_t,$$
$$f_t^n = A_{f,n} + B_{f,n}' X_t. \tag{12}$$

Coefficients A_f and B_f are such that the interest rate is

$$i_t = \delta_0 + \delta_1' X_t, \tag{13}$$

and the risk premium for the n-maturity bond is

$$E_t rx_{t+1}^n = -(1/2) var_t(rx_{t+1}^n) + cov_t(rx_{t+1}^n, e_{t+1}')\lambda_t$$
$$= \left[-(1/2)B_{n-1}' \Sigma B_{n-1} + B_{n-1}' \Sigma \lambda_0\right] + \left[B_{n-1}' \Sigma \lambda_1\right] X_t. \tag{14}$$

Define e_t^* by

$$X_t = \mu^* + \Phi^* X_{t-1} + e_t^*, \qquad \text{where } \mu^* = -\Sigma\lambda_0,$$
$$\Phi^* = \Phi - \Sigma\lambda_1. \tag{15}$$

There exists a risk-neutral probability measure under which $e_t^* \sim \mathcal{N}(0, \Sigma)$ and is independent over time. In the solution,

$$f_t^n = -(1/2)B_{n-1}' \Sigma B_{n-1} + E_t^* i_{t+n-1}, \tag{16}$$

where operator E^* takes expected value using the risk-neutral probability measure.

Given μ^* and Φ^*, coefficients A, A_f, B and B_f do not depend on λ_0, λ_1, μ or Φ.

Proof. See the appendix. □

Equation (14) describes bond risk premia according to the model. We interpret the covariance term as the amount of "risk" of each bond - how they correlate with existing shocks. Therefore, we interpret λ_t as the market "price" of such risks. For example, $\lambda_t(3)$ (the third element of λ_t) is the price of shocks to the term structure slope factor.

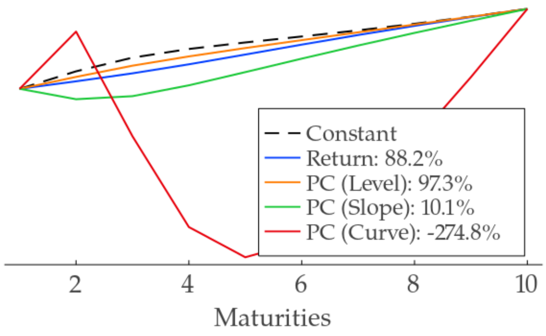
As the second line of (14) shows, the two variance terms are constant over time - the model is conditionally homoscedastic -, and one can thus drop the t subscript. Variation in risk premia $E_t rx_{t+1}$ comes from λ_t, only. So, to replicate (10), all we need is to pick the right λ_0 and λ_1.

The second part of proposition 1 describes the model's solution in terms of the risk-neutral dynamics. Equation (16) shows that, under the risk-neutral measure, the expectations hypothesis holds, up to a Jensen's inequality additive constant. Forward rate variation reflects *only* variation in expected interest, not variation in term premia (recall (1)).

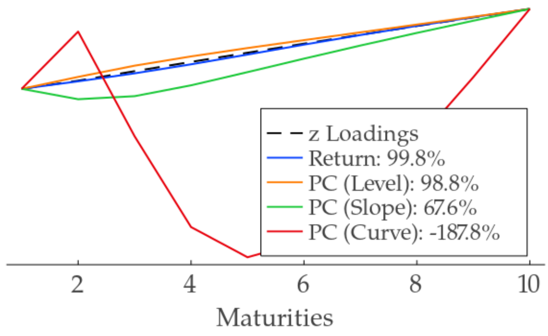
4.2. Estimation Procedure and Zero Restrictions

One way to estimate the model is to run OLS on (9), and then compute λ_0 and λ_1 by matching (10) and (14). If we allow the three principal component factors in X_t to predict returns along with z, so that the restrictions imposed by (10) no longer hold, Adrian et al. (2013) show that this procedure prices the term structure with high precision.[1] The downside is that, first, we rely on many free parameters (λ_0

[1] Equation (10) is a restricted version of the more general case $rx_t^n = \alpha_n + b_n X_t + \epsilon_{n,t+1}$.



(a) Dashed = Constant α_n ($+(1/2)\text{var}(rx_t^n)$)



(b) Dashed = z Loading β_n

Notes: I estimate (9) by OLS and compute residuals $\hat{e}$. Each panel shows the covariance between excess returns rx_t^n and factor i innovations $\hat{e}_{i,t}$. Each solid curve corresponds to a different factor, and the x-axis contains bond maturity. The dashed curves plot the empirical constant term α (panel 3a, added to $(1/2)\text{var}(rx_t^n)$) and loading β (panel 3b) from the empirical model (10). I scale the curves so that they take the same value for $n = 10$.

Figure 3: Cross-Sectional Amount of Risk *vs* Empirical Model (10)

and λ_1 contain a total of twenty parameters in a four-factor model). Second, the estimation is sensitive to model specification - changing (9) leads to wildly different results - and sampling uncertainty.[1]

I instead follow Cochrane and Piazzesi (2009) and estimate factors' (risk-neutral) dynamics solely from the cross section of yields. We can then use the model's restrictions to infer the true dynamics.

More specifically, note from (13) and (16) that we can write $f_t^n = \delta_1' \Phi^{*n-1} X_t$ plus a constant. So, given data for f and X, we can estimate Φ^* (and μ^*, which enters the constant). This procedure explores the fact that prices depend only on the risk-neutral dynamics (μ^*, Φ^*), a point highlighted by the last claim in proposition 1.

Price of Risk Variation. The step after that is to estimate λ_1, which then implies the data-generating dynamics via (15). To do it, I match (14) and (10): I force the affine model to match the empirical return-forecasting regression. Its main restriction is that loadings on the level, slope and curvature factors must be zero: *the price of risk varies only due to variation on the return-forecasting factor*. Given the theoretical loadings $[B_{n-1}' \Sigma \lambda_1]$, I satisfy the restriction if the corresponding *columns* of λ_1 are zero. In that case, (14) becomes $E_t rx_{t+1}^n = [B_{n-1}' \Sigma \lambda_1(1)] z_t$ plus a constant, where $\lambda_1(1)$ is the first column of λ_1. We reduce the twenty parameters in (λ_0, λ_1) to eight.

Priced Shocks. I take one step further and also assume that only a single shock is priced. In economics language, that is to assume that marginal utility is orthogonal to the remaining three shocks. In the model, it means all but one entry of vector λ_t are equal to zero.

It turns out that we do not need more than one priced factor to reproduce the empirical regression (10), as I show in figure 3. I start by running OLS on (9) and computing residuals $\hat{e}_t$ (the appendix shows the series). For each model factor i, I then plot a solid-line curve that depicts the covariance between its innovations $\hat{e}_{i,t}$ and excess returns on bonds of different maturities rx_t^n. They represent the estimated amount of factor-i risk of each bond. The dashed lines represent parameters α (plus the Jensen's inequality term $(1/2)\text{var}(rx^n)$, panel 3a) and β (panel 3b) from equation (10). I estimate both by OLS. The horizontal axis contains bond maturity n. Since $rx_t^1 = 0$, all curves start at zero. I scale them so the plot looks nice - we are interested only in their shape, not their scale or sign.

Given the estimation strategy of matching (10) and (14), and the assumption that $\lambda_t = \lambda_0 + \lambda_1(1) z_t$, our job is to find linear combinations of the solid-line curves that match the dashed-line ones; λ_0 and $\lambda_1(1)$ are the weights. The figure makes it clear that a single curve will do the job, in both cases. In the

[1] For example, in the case of highly persistent processes, OLS estimates lead to a downward bias on the system's persistence (Yamamoto and Kunitomo (1984)). This is particularly problematic in applications for term structure pricing, since risk-neutral dynamics must have a near-unitary root (the level factor alone explains about 90% of forward rate movements). Indeed, several bias-correction methods have been proposed to address this issue. See Phillips and Yu (2005), Tang and Chen (2009) and Bauer et al. (2012).

legend, I report the R^2 from regressing each curve individually on α (+ Jensen's inequality term) and β. Cross-sectional covariances of excess returns with both return factor shocks and level factor shocks provide, by themselves, an almost perfect fit to the empirical model (14).[1] So, I assume throughout that only a single shock is priced. That is to say that only one element of λ_t is different from zero; the same is thus true for λ_0 and $\lambda_1(1)$. I therefore further reduce the number of estimated parameters in (λ_0, λ_1) to two.

Estimation. I estimate δ_0, δ_1, μ^*, Φ^* by minimizing model pricing errors. Given a choice for priced factor $i \in \{1, 2, 3, 4\}$, note that $\mu^* = \Sigma(i)\lambda_0(i)$, where $\Sigma(i)$ is the i-th row of Σ, and $\lambda_0(i)$ is the i-th element of λ_0. I therefore search on the space of $\lambda_0(i)$ parameters. Covariance matrix Σ I estimate as $\hat{\Sigma} = \sum_t \hat{e}_t \hat{e}_t' / (T-1)$.

In all, let $\theta = (\delta_0, \delta_1, \lambda_0(i), \Phi^*)$. The problem is to minimize the mean squared error implied by the model:

$$\text{Min}_\theta \left\{ \frac{1}{N} \sum_{n=1}^{N} \left[\frac{1}{T} \sum_{t=1}^{T} \left(f_t^n - A_{f,n} - B_{f,n}' X_t \right)^2 \right] \right\}^{\frac{1}{2}}$$

I write the objective function above so as to highlight that this is a method-of-moments estimation. The model contains the stochastic singularity $\eta_{n,t} = f_t^n - A_{f,n} - B_{f,n}' X_t = 0$ which implies $E(\eta_{n,t}^2) = 0$ for all n. The optimization above simply minimizes an equally-weighted combination of these moment conditions.[2] It is not quite a GMM procedure since I do not minimize a quadratic form. If that was the case, and assuming an identity weighting matrix, the brackets term of the objective function above would be squared.

I solve the problem using numerical methods. Having estimated θ, we can calculate coefficients B_n. In the model's solution, $E_t rx_{t+1}^n = [B_{n-1}' \hat{\Sigma}(i) \lambda_1(1, i)] z_t$ plus a constant, where $\lambda_1(1, i)$ is the element $(1, i)$ of matrix λ_1 - the last parameter to be estimated. To match the loadings of the empirical return model (14), I run the cross-sectional regression

$$\beta_n = \lambda_1(1, i)(B_{n-1}' \hat{\Sigma}(i)) + \text{error}, \tag{17}$$

in which $B_{n-1}' \hat{\Sigma}(i)$ is the right-hand variable. Having found λ_1, we can infer the data generating dynamics (μ and Φ) using (15). This completes the estimation.

4.3. Estimation Results

Table 4 contains most results. In Panel A, I compare models and estimation strategies. For each case, the table shows the largest eigenvalue of matrices Φ^* and Φ, the mean squared error in bond pricing and the total number of free parameters.[3] Panel B reports the estimated parameters in the baseline specification.

In this baseline, I assume X contains four factors and that *only level factor shocks are priced*. The table shows that this choice leads to the best model fit, and I also support it under the evidence (presented later) that the level factor is associated to inflation and US risk premium. There is a total of 33 estimated parameters.

The exponential affine model yields a mean squared error of 0.22% and a mean absolute error (not shown in the table) of 0.18%. I compare that to two benchmarks. The first is the unrestricted OLS model $f_{n,t} = a_n + b_n X_t + e_t$, which yields the lowest mean squared errors (0.05%) in the class of linear models at the cost of 50 free parameters (ten equations, five coefficients each), plus 16 if we estimate the state's law of motion (9).

[1] In the US case, Cochrane and Piazzesi (2009) show that only covariances with level-factor shocks approximate empirical loadings on the return forecasting factor. In their case, the amount of risk on the other shocks has completely different shapes in the cross-section. Sadly, that is not the case in Brazil.

[2] Due to the stochastic singularity, maximum likelihood is not feasible - one would have to assume measurement errors in the model.

[3] Let $\hat{f}_{n,t}$ be the forward rate predicted by a model. The mean squared error is $\left\{ \frac{1}{N} \sum_n \left[\frac{1}{T} \sum_t (f_{n,t} - \hat{f}_{n,t})^2 \right] \right\}^{\frac{1}{2}}$. That is the objective function I minimize.

Panel A. Stability and Model Fit					
Model	$\|\|\text{eig}(\Phi^*)\|\|_\infty$	$\|\|\text{eig}(\Phi)\|\|_\infty$	MSE (%)	$[\Phi - \Phi^*]_{2,1}$	# Param
Benchmarks					
OLS			0.05		50 (+16)
VAR	1.05	0.40	0.07	-1.3	55
Baseline	1.01	0.84	0.22	-1.4	33
3 factors	0.99	0.75	0.31	-1.4	21
2 factors	0.99	0.71	0.51	-1.3	12
Free μ	0.99	0.83	0.06	-1.4	37
$t < 2020$	1.00	0.94	0.25	-2.2	33
Return factor priced	1.03	0.89	0.29	-1.3	33

Panel B. Estimated Parameters (Baseline)					
	Constant	Return z	Level	Slope	Curve
	Risk-Neutral Dynamics μ^*, Φ^*				
Return z	0	0.61	0.23	-0.08	0.57
Level	0.01	-0.05	0.93	0.50	-0.12
Slope	0	0.05	0.05	0.40	0.80
Curve	0	-0.19	0.11	0.16	0.69
	True Dynamics μ, Φ				
Return z	0	0.11	0.23	-0.08	0.57
Level	0	-1.42	0.93	0.50	-0.12
Slope	0	-0.03	0.05	0.40	0.80
Curve	0	-0.19	0.11	0.16	0.69
δ	0.10	0.03	0.48	-0.81	0.27

Notes: Panel A compares different model performance under different estimation methods and underlying assumptions. The $||\text{eig}(\Phi^*)||_\infty$ column indicates the largest eigenvalue matrix Φ^*, and the analogous is true for $||\text{eig}(\Phi)||_\infty$. MSE is the forward rate mean squared error and MAE the mean absolute error; # Param indicates the number of free parameters in each model/estimation strategy. Panel B reports the estimated parameters μ, Φ, μ^*, Φ^*, δ_0 and δ_1 in the baseline estimation.

Table 4: Estimation Results and Model Dynamics

As a second benchmark, I follow a procedure similar to Adrian et al. (2013) to estimate the model. It involves estimating Φ from factor data and (λ_0, λ_1) from cross-sectional regressions of excess returns on sample innovations. The appendix provides details. Unlike unrestricted OLS, this procedure involves active zero restrictions, fewer parameters and a good fit that is almost as good (0.07%).

In our baseline case, the largest eigenvalue of Φ^* is 1.01, which is close to a unitary root. This is an expected result, which we verify in all specifications. Under the risk-neutral measure, there is no term premium (except for a Jensen's inequality term): forward rates equal expected future interest rate (equation (16)). Since forward rates of different maturities tend to move in block - the level principal component captures most variation in the term structure - innovations have to be long-lived so that they affect interest rates far in the future. Hence, the near-unitary root.

Panel B of table 4 reports the estimated parameters in the baseline exercise. By inspecting Φ^*, we see that persistence follows mostly from the level factor, with an autoregressive coefficient of 0.93. The estimate of δ_1 shows that the level factor also impacts interest rate the most (δ_1's loading on the slope factor -0.81 is about twice the loading on the level factor, but the level factor's standard deviation is almost four times as large as that of the slope factor).

Estimation of λ_1 - its single entry - then yields an estimate of the actual dynamics Φ via (15). Only the first column of λ_1 is non-zero, so only the first columns of Φ^* and Φ differ. The table shows that, in fact, only the first two rows of their first column differ by more than two decimal digits. Under the true probability measure, the autoregressive coefficient of the return factor declines from 0.61 to 0.11, and its expected effect over next-period level factor goes from -0.05 to -1.42.

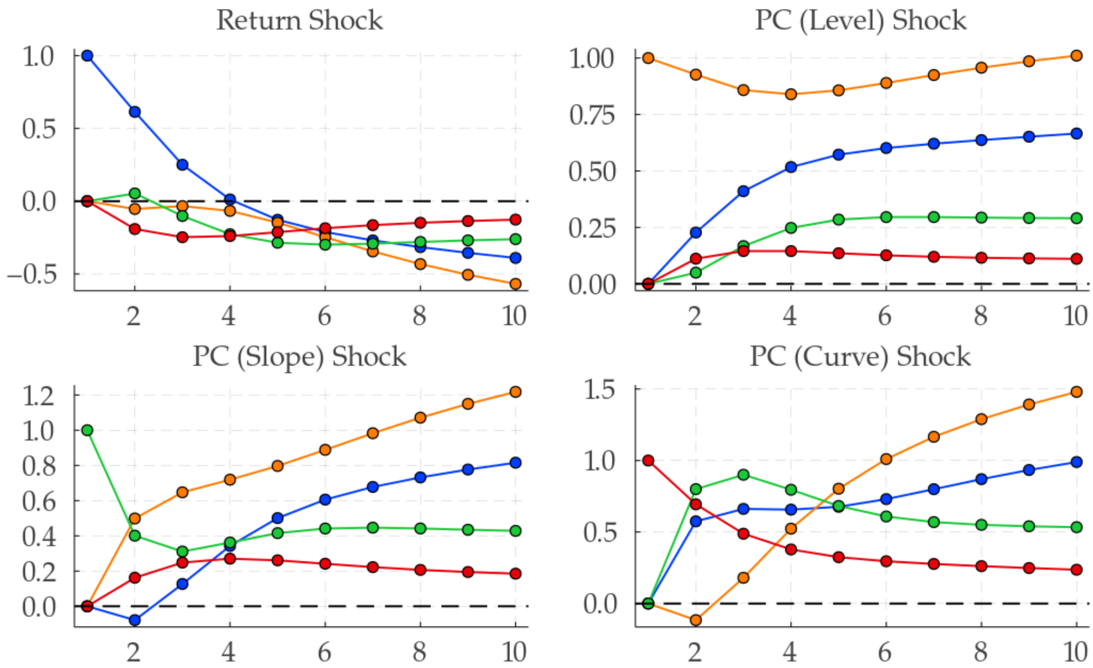
Figures 4 and 5 plot the impulse response functions of the system under the risk-neutral and actual probability measure, respectively. They make it easier to understand the dynamics Φ^* and Φ represent. See the figures' notes for the color legend.

Figure 4 exposes the system's near-unitary and in fact explosive root under the risk-neutral probability measure. An increase in the level factor (top, right panel) leads to an immediate and prolonged increase in all the other factors. The fact that the return factor increases is particularly important. The same observation is true when we look at slope and curve shocks (bottom panels): both lead to an increase in all variables in the system. The exception here is the impulse to the return factor (top left panel), which leads to a *small*, yet protracted decline in the level, curve and slope factors. In all cases, movements in the level factor are long-lived, and interest rates are thus persistent under the risk-neutral measure.

Dynamics under the true probability measure (figure 5) have one critical difference: *an increase in the return factor leads to a larger decline in the level factor*. This is where the -0.05 to -1.42 change in the Φ matrices shows up. Large bond risk premia forecast lower yields for bonds of all maturities. That, of course, leads to changes in the impulse response to the three other shocks as well. Under the risk-neutral measure, they lead factors to move in the same direction, as we saw. Under the true measure, however, the return factor "pushes" the level factor in the opposite direction. *That stabilizes the system's dynamics*. The largest eigenvalue of Φ drops from 1.01 to 0.84. Part of that change must also be attributed to fact that the return factor has a lower autocovariance coefficient under the true measure (0.11 vs 0.61), which reduces the amplitude of the cycles that characterize the IRFs.

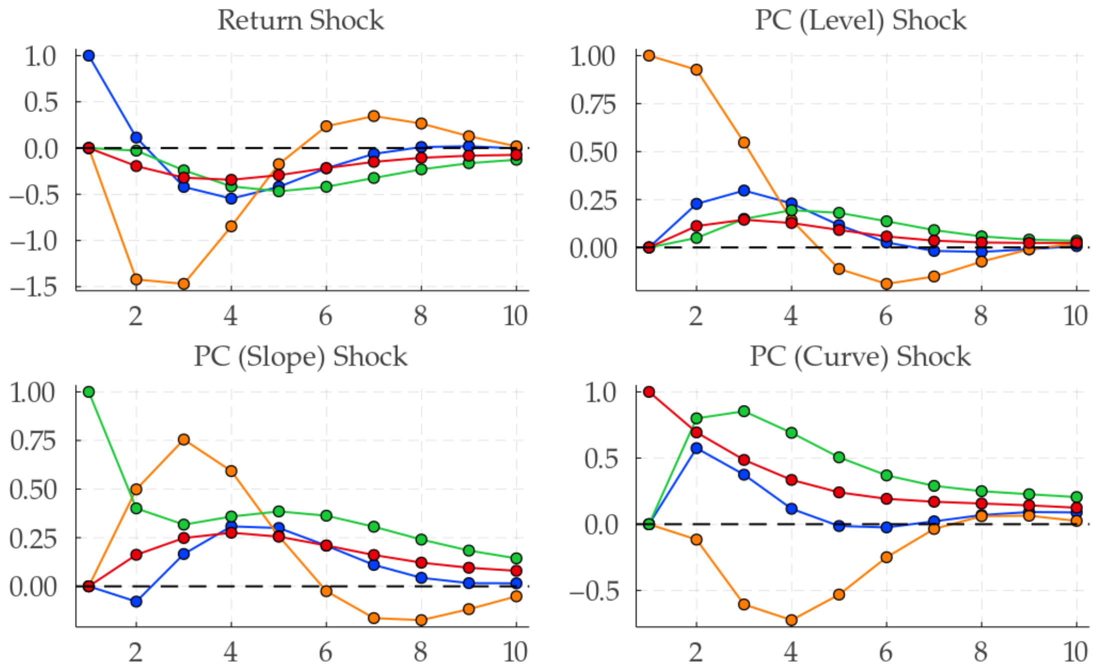
The interaction between return and level is particularly visible in the impulse response to a level shock. As yields increase in period one, it is still true that risk premia increase relative to average for a few periods (about 5 years). But now, such increase causes a decline in yields - the orange curve goes below zero five periods after the initial shock. The interaction goes on: a lower level factor brings down the risk premia to the point that its decay is reversed once again until the system stabilizes.

We can verify this periodic sign reversal of the level factor in the response function to the other shocks too (although somewhat less pronounced in the case of curve shock). This is good news. It means that the model can account for the swings of the Brazilian interest rates verified in sample (red curve in plot 1a). High interest today forecasts low interest four to five years ahead. Analytically, this feature of the model follows from the negative difference between the entry (2,1) of matrices Φ and Φ^*, which I report for all specifications in Panel A of table 4.



Notes: Impulse response function of model dynamics under the risk neutral probability measure. Blue = return factor. Orange = level factor. Green = Slope factor. Red = Curve factor.

Figure 4: Risk-Neutral Dynamics Φ^*



Notes: Impulse response function of model dynamics under the true probability measure. Blue = return factor. Orange = level factor. Green = Slope factor. Red = Curve factor.

Figure 5: Data-Generating Dynamics Φ

Where does it come from? By assumption, level shocks directly affect marginal utility, and since bond prices go down when the level factor jumps up, we estimate $\lambda_1(2,1) < 0$. (You can see this from equation (17), $\beta_n > 0$ and $B_{n-1} < 0$ in the second column.) From the second line of (15), $\lambda_1(2,1) < 0$ leads to $\Phi_{2,1} < \Phi^*_{2,1}$.

Intuitively, we need the sensitivity of marginal utility to level shocks to depend on the return-forecasting factor state - such heteroskedasticity of the stochastic discounting is what allows for time-varying bond premia. When the return factor is positive, marginal utility is more sensitive to level shocks, and investors demand a higher return to hold bonds. Risk-neutral pricing internalizes that by increasing the chance of a higher level shock in the following period: $\Phi_{2,1} < \Phi^*_{2,1}$.

To sum up, the dynamics of state variables are characterized by the following key patterns. Under the risk-neutral measure, level factor movements "pull" the return factor in the same direction; shocks are long-lived. Under the actual measure, the level factor "pushes" the level factor to the opposite direction, leads to periodic swings in the interest rate and stabilizes the system.

Slope, Curve and Fama-Bliss Revisited. Dynamics of the slope and curve factors are important too. For example, the system's response to a slope shock reconciles the results of the Fama-Bliss regressions (table 1), which partially supported the expectations hypothesis, with return predictability. The slope factor is responsible for most movements in forward spreads $f^n_t - i_t$. The IRFs show that, following a positive slope shock, which increases spreads, the level factor increases, and so do interest rates. Indeed, column "Model β_n" of table 1 reports the model-implied coefficients from the Fama-Bliss regressions. All of them are within one and a half standard deviations from their empirical counterparts.

I also calculate the model-implied coefficients from the twin regression (3), of excess returns on term spreads. The model predicts that the two and three-year term spreads do not predict excess returns. For larger maturities, it does but with coefficients that are still hard to differentiate from zero. Yet, there is predictability. We *can* forecast returns, but the term spread is a poor choice for a predictor. We need the tent-shaped profile of coefficients, not spreads.

Finally, shocks to the curve factor are less volatile, but have the interesting property of being the only shock in the model that sends the return factor and the level factor on opposite directions.

Alternative Specifications. Panel A of table 4 summarizes results from various different specifications of the model. The "3 factors" and "2 factors" rows correspond to specifications with less than four state variables (I cut the curve factor and then the slope factor). Mean squared errors naturally get larger.

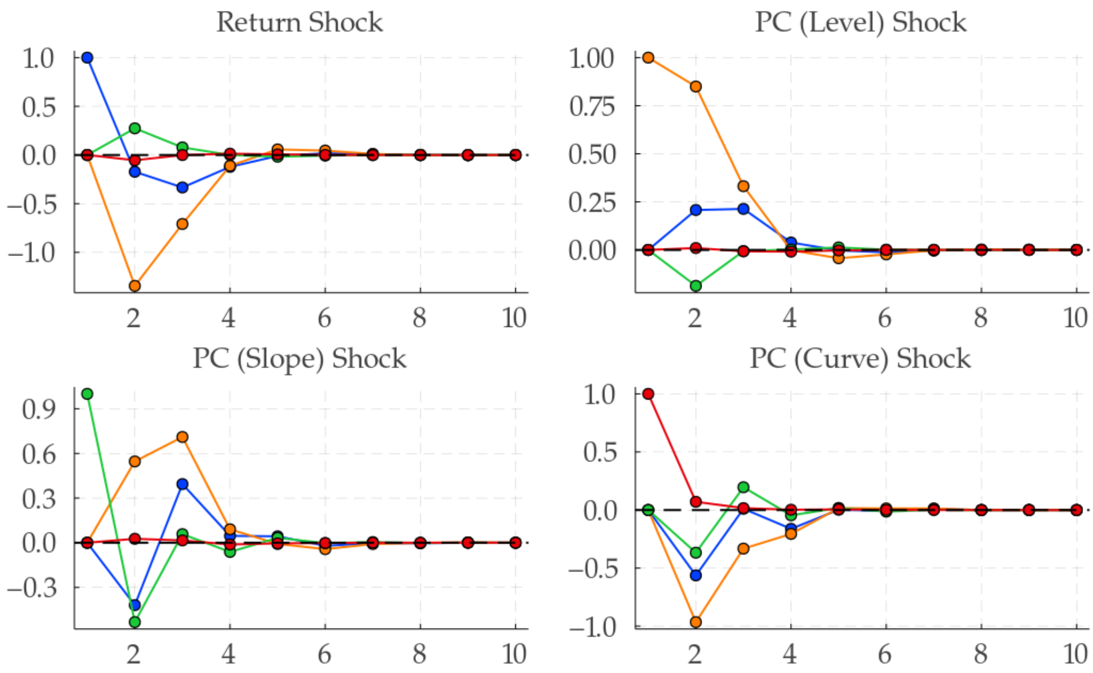
In the "free μ" case, I add a constant μ to the law-of-motion of (9), which becomes $X_t = \mu + \Phi X_{t-1} + e_t$, and adjust (15) accordingly: $\mu^* = \mu - \Sigma\lambda_0$. The change relaxes the assumption that the sample average of X equals its unconditional average. At the cost of four additional parameters, the model yields mean squared errors smaller than that of OLS - the best linear predictor - by less than 0.02%. This shows that the exponential-affine model *can* lead to small pricing errors.

The "Return factor priced" row considers the case in which not level factor shocks but return factor shocks are priced. Model fit does is slightly worse. The poorer performance in model fit is reminiscent of the fact that the cross-sectional covariances between bond returns with the return factor shocks do not provide the best fit to the empirical excess return model (10).

The different specifications yield a few common results. Risk-neutral dynamics have a root close to unity, while under the true-probability measure dynamics are less persistent (as measured by the largest eigenvalue of Φ). More importantly, our characterization of the dependency of the return and level factors on each other's lagged values is also robust, as evidenced by the difference between entry (2,1) of Φ and Φ^*, negative in all cases.

Figure 6 provides one example. In it, I plot the impulse-response function of the model estimated from the VAR (9) with a constant μ. The VAR estimates a model with lower persistence (0.40 *vs* 0.84 in the baseline), but qualitatively the responses of return and level shocks to each other are similar: a high level factor increases the return factor; a higher return factor lowers the term structure level. This is an encouraging result. The key dynamic property of the model, which I estimate without the use of time series autocovariances, also shows up in time series autocovariances.

In the appendix, I provide the impulse-response functions to the other specifications as all. These



Notes: Impulse response function of model dynamics under the true probability measure. Blue = return factor. Orange = level factor. Green = Slope factor. Red = Curve factor. IRF estimated from direct estimation of $X_t = \mu + \Phi X_{t-1} + e_t$.

Figure 6: Data-Generating Dynamics Φ - Estimated from VAR

two properties are present in all of them.

Interpretation. How can one interpret the different behavior of our state variables under the actual and risk-neutral measures? "Risk-neutral" measure is another name to probability times marginal utility. In factoring in the latter, one can do asset pricing by simply taking expectations of uncertain payoffs. Hence, the economic interpretation of an asset pricing model comes entirely from the discount factor.

The exponential-affine model offers a lot of tractability to see this point. In a given period $t+1$, both the actual shock e_{t+1} and the artificial risk-neutral "shock" e^*_{t+1} are Gaussian with the same covariance matrix Σ. Their probability density functions therefore have the same normalizing constant $\mathcal{M}$; only their exponential kernel differ. Hence, we can factor in the discount factor m_{t+1} by changing such exponential kernel and put into evidence the density function of the risk-neutral measure:

$$1 = E_t \exp\{m_{t+1}\} R_{t+1}(e_{t+1}) = \mathcal{M} \int \exp\{-(1/2)e'\Sigma^{-1}e\} \exp\{-i_t - \lambda_t'\Sigma\lambda_t - \lambda_t e\} R_{t+1}(e) de$$
$$= \mathcal{M} \int \exp\{-(1/2)(e-\phi_t)'\Sigma^{-1}(e-\phi_t)\} \exp\{-i_t\} R_{t+1}(e) de.$$

So, risk-neutral pricing here simply amounts to taking expected value of returns "as if" shocks e_{t+1} had a conditional mean equal to ϕ_t. The last equality then implies

$$\phi_t = -\Sigma\lambda_t,$$

and all we need now is to interpret ϕ_t. To that end, suppose period $t+1$ comes and we observe $e_{t+1} = \phi_t$. State variables will then be given by

$$X_{t+1} = \Phi X_t + \phi_t = \Phi X_t - \Sigma\lambda_t = \mu^* + \Phi^* X_t.$$

We have thus a link between the estimated risk-neutral dynamics - to which the IRFs above give a useful interpretation - and the discount factor m. What the equation above tells us is that *the discount factor, or marginal utility, places elevated weight on combinations of shocks that lead our states X to present the*

Right-Hand Variable	Return Factor	Level Factor	Slope Factor	Curve Factor
International Variables				
US Term Spread	0.04	0.35	0.00	0.21
MSCI	0.21	0.19	0.00	0.02
EMBI (Brazilian bonds)	0.12	0.05	0.03	0.16
Δ Commodities' Prices (CRB)	0.03	0.00	0.00	0.02
Domestic Variables				
GDP growth	0.00	0.01	0.03	0.03
Short-term interest	0.19	0.83	0.17	0.00
Inflation	0.02	0.38	0.13	0.11
Δ Industrial Production	0.01	0.00	0.00	0.12
Trade Balance	0.03	0.03	0.06	0.01
Current Account	0.02	0.02	0.12	0.00
Primary Surplus	0.06	0.23	0.06	0.00

Notes: I run regressions of the form $x_t = a + bm_t + \varepsilon_t$, where x is a model factor and m is a macroeconomic variable. The table reports the R^2 of each regression.

Table 5: R^2 of Regressions of Model Factors into Macro Variables

risk-neutral dynamics (μ^*, Φ^*), *such as* ϕ_t.

Under the risk-neutral dynamics, the level factor increases 0.01 per period unconditionally (that is the μ^*). Since factors have zero mean, this indicates that, on average, marginal utility increases in the level factor, which is also implied by (and equivalent to) $\lambda_0(2) < 0$. Marginal utility tends to be higher in times of high bond yields, such as the 2015-16 recession.

Conditional on the state X, under the risk-neutral probability measure, the system takes longer to return to its long-term average. Given a, say, high return factor, marginal utility concentrates on states of nature in which the return factor stays high and the level factor fails to decline. By the argument above, investors fear (or marginal utility growth is high) in shocks that prolong the period of high risk premium and bond yields.

All these observations ask for an investigation of which macroeconomic variables are correlated with the level and return factors. To investigate correlation, I report in table 5 the R^2s of regressions of the form

$$x_t = \alpha + \beta\, m_t + \epsilon_t$$

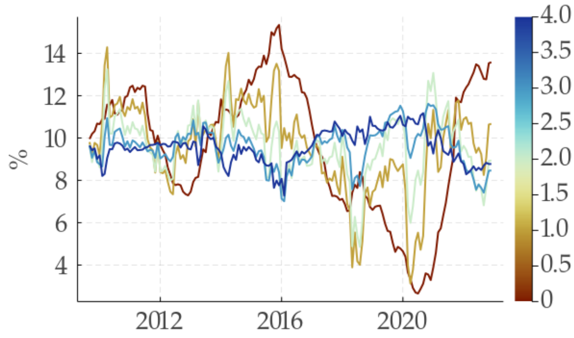
where x is one of my factors and m is a macroeconomic variable of choice.[1] Given obvious endogeneity issues, I do not report point estimates. I am interested only in the correlation between variables, not causal or structural relationships.

The MSCI index and short-term interest rates offer a interesting fit for the return factor, which is disconnected from all other macroeconomic variables I put on the right-hand side.

The level factor, on the other hand, is highly explained by the interest rate (as expected), the US term spread and domestic inflation, with R^2s of 0.83, 0.35 and 0.38 respectively. The correlation with primary surplus (R^2 = 0.23) is also high, probably as a result of the decline in bond prices with the larger deficits in 2015-16. Therefore, bond yields tend to be larger in times of domestic interest and inflation, negative fiscal news, or in times of higher US risk premium, as proxied by the term spread (see Fama and Bliss (1987)).

This result makes the story told by the model more interesting and strengthens the case for a price level factor shock. Innovations to the term structure level affect marginal utility because they usually translate news about inflation, fiscal policy or foreign risk premium. Given a large domestic premium,

[1] Inflation is the annual growth in the consumer price index. GDP is the log real gross domestic product. Industrial production is the log of the quantum index of industrial production. The primary surplus, trade balance and current account are expressed as ratios of GDP and cumulated in the last twelve months. The proxy for commodities' prices is the Commodity Research Bureau BLS Spot index, taken from Bloomberg. MSCI is a free-float weighted equity index that captures large and mid-size stocks across Emerging Market countries. I take from Bloomberg.



(a) Expected Interest $E_t i_{t+n-1}$

(b) Term Premium $f_t^n - E_t i_{t+n-1}$

Notes: I use the baseline estimation of the exponential-affine model to decompose forward rates into expected interest and term premium, as shown by equation (1). Each curve corresponds to a different maturity, and the colors match the terms of the same decomposition across figures. For instance, the blue curves represent $E_t i_{t+4}$ on the left, and Term Premium$_t^5$ on the right; and $E_t i_{t+4}$ + Term Premium$_t^5 = f_t^5$.

Figure 7: Decomposition of the Term Structure of Forward Rates

which typically accompanies a high level state given the positive correlation between the two shocks, investors care about innovations that prolong the scenario of large levels, that is, large domestic interest and inflation, or international risk premium.

Finally, I do not find large R^2s on the regressions with the slope and curve states.

Yields Decomposition. The exponential-affine model offers a decomposition of the yield and forward-rate curves. I focus on the latter. From (13) we have

$$E_t i_{t+n-1} = \delta_0 + \delta_1' E_t X_{t+n-1}.$$

From (9), we have expected factors $E_t X_{t+n-1}$. With $E_t i_{t+n-1}$, we can use (1) to infer the term premium or directly calculate it from its formula in (1).

Figure 7 contains the yield curve decomposition. Plot 7a depicts expected interest rates, plot 7b depicts term premia. Added up, they yield the model's forward rates. The format of the graphs is similar to 1a: different colors correspond to different bond maturities, as indicated by the bar on the right. I choose maturities so that the terms summing up to the same forward rate have the same color.

While we estimate that the economy's dynamics are highly persistent - the largest eigenvalue of Φ is 0.84 in an annual frequency -, the pattern of expected interest rates does not resemble the random walk $i_t = i_{t-1} + e_t$, as Cochrane and Piazzesi (2009) find for the US. High interest rates today are usually associated with lower interest rates three or four years in the future.

Unintuitive as it is, visual inspection of the data does not reject this result. Interest in Brazil (red curve in figure 1a) has clearly moved in waves during the 2010s, with successive up-and-down swings. Forecasting that it will continue to do so is a sensible prediction. It follows from the combined dynamics of the return and level factors, which I show to be robust across model specifications. Any shock that reduces risk-premia (z goes down) leads to higher yields (and interest rate); higher yields lead to an increase in risk premia, which in turn reduces yields, and so on.

The swings do not deny the conclusion that shocks are long-lived. The fact that the economy takes time to return to its long-term average does not mean we should expect a given shock to have similar sign effects on a certain variable in different time horizons.[1] In fact, if I plotted maturities above $t + 4$ (I do not, to avoid making the plots too crowded), we would see that expected future interest would *not* continue to move in the opposite direction of current interest. It would instead fluctuate with up and down movements for a long time until it stabilized around its long-term average.

Plot 7b of term premia also reveals interesting patterns. First, the term premium is non-zero and economically significant. At the height of the 2015-16 crisis, it reaches 7.5% for some bonds. More

[1]Mathematically, matrix Φ contains a couple of complex eigenvalues that lead to wave movements that characterize the dynamics of X.

importantly, term premia of different maturities do not move in lockstep. The premium does not necessarily increase with maturity always. As equation (1) suggests, the underlying reason is the projected path of excess returns, which varies depending on the state of the models' factors X.[1]

For instance, the two years that feature the more pronounced surges in term premia (on average) are 2015 and 2018. On the former, term premia increase from maturity $n = 2$ to $n = 5$. On the latter, it decreases.

It turns out that, in this example, we can understand the different behavior of term premia only from the joint dynamics of the return and the level factors I highlight above. As I show in the appendix, in 2015 both peak in (by then) all-time highs. 2018, on the other hand, features an again high return factor (even higher than it was in 2015), but a level factor that is close to and then lower than the long-term average of zero.

As the impulse-response function figure 5 indicates, in case of a pure return factor shock (top, left plot), agents expect the return factor itself initially to increase, of course, and then decline to negative territory starting in period three. That means that they expect larger excess returns up to period three (remember: z_t predicts *one-year ahead* returns) and then *lower* excess returns. Thus, the terms of the sum that account for the term premia are positive up to maturity three, and become negative for larger ones.[2] Hence the increasing term premia up to period three and their decline for larger maturities.

But in 2015 the level factor was also unusually high. The model is linear, so we can try to understand the profile of term premia by summing up the system's responses to a return factor shock and to a level factor shock. The latter is the novelty now. As concluded before (and shown by the top-right graph of figure 5), an above-average level factor induces an increase in the return factor z *for a few years in the future*. Such increase, combined with the effects of the other two model factors, to a large extent offsets the lower mid-to-long term risk-premium that follows the innovation in the return factor, as explained in the paragraph above. Unlike in 2018, in 2015 agents did not expect risk premium to quickly revert to average. Its drawn-out path accounts for a cross section of term premium that increases in maturity. Precisely what we see in figure 7b.

5. Final Remarks

This paper studies bond risk premia in Brazil using twelve years of yield data, including that of long-term bonds. I find that some empirical patterns found for the US case are also present in the Brazilian case (like return predictability through the Cochrane and Piazzesi (2005) tent-shaped combination of forward rates), and others are not (like the result from Fama-Bliss regressions). I find that the return factor that predicts bond excess returns also forecasts stock returns and returns on the carry trade, which enlarges its interpretation to be of a return factor for Brazilian asset markets (at these three important markets, at least). Further studies that attempt to find a common component to the risk premium of different classes of assets would be welcome. As would papers considering whether such common component also affects risk premia in other emerging markets.

I estimate an exponential-affine model that characterizes the joint distribution of the return factor and the term structure's principal components. Even with more zero restrictions than researchers usually assume, the model can price the term structure almost as well as unrestricted OLS. It also reconciles the result from the Fama-Bliss regressions with predictability.

Different specifications of the model and of the estimation procedure lead to a common property of the joint dynamics of states under the true probability measure. Surges in the level factor forecast an increase in risk premia, and an increase in risk premia reduces bond yield levels. The latter effect does not hold under a risk-neutral measure.

To the extent that level shocks affect marginal utility and typically translate higher domestic inflation or US risk premia, one can interpret this difference in dynamics as investors' fear, and hence pricing, of sequences of innovations that prolong periods of high bond yields despite the elevated premia. Intuitive

[1]This is what Cochrane and Piazzesi (2009) call a non-trivial term structure of risk premia.

[2]The model implies that the betas in (10) are positive and increasing, just like q_n in table 2. Therefore, the terms $E_t(rx^n_{t+1} - rx^{n-1}_{t+1})$ in the definition of the term premium always have the same sign as z.

as it is, I do not however present a consumption model that provides a structural, macroeconomic explanation to this robust finding. I think such effort constitutes an interesting and necessary venue for future research.

References

Adrian, T., Crump, R. K., and Moench, E. (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics*, 110(1):110–138.

Aguiar, M. and Gopinath, G. (2006). Defaultable debt, interest rates and the current account. *Journal of International Economics*, 69(1):64–83.

Aguiar, M. and Gopinath, G. (2007). Emerging Market Business Cycles: The Cycle Is the Trend. *Journal of Political Economy*, 115(1):69–102.

Akgiray, V., Baronyan, S., Sener, E., and Yılmaz, O. (2016). Predictability of Emerging Market Local Currency Bond Risk Premia. *Emerging Markets Finance and Trade*, 52(7):1627–1646.

Almeida, C., Ardison, K., and Kubudi, D. (2015). Approximating Risk Premium on a Parametric Arbitrage-Free Term Structure Model. SSRN Scholarly Paper ID 2535766, Social Science Research Network, Rochester, NY.

Almeida, C. and Faria, A. (2014). Forecasting the Brazilian Term Structure Using Macroeconomic Factors. *Brazilian Review of Econometrics*, 34(1):45–77.

Ang, A., Dong, S., and Piazzesi, M. (2007). No-Arbitrage Taylor Rules. Working Paper 13448, National Bureau of Economic Research.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–787.

Bauer, M. D., Rudebusch, G. D., and Wu, J. C. (2012). Correcting Estimation Bias in Dynamic Term Structure Models. *Journal of Business & Economic Statistics*, 30(3):454–467.

Bekaert, G. and Hodrick, R. J. (2001). Expectations Hypotheses Tests. *The Journal of Finance*, 56(4):1357–1394.

Boz, E., Daude, C., and Bora Durdu, C. (2011). Emerging market business cycles: Learning about the trend. *Journal of Monetary Economics*, 58(6):616–631.

Calvo, G. A., Izquierdo, A., and Talvi, E. (2006). Sudden Stops and Phoenix Miracles in Emerging Markets. *The American Economic Review*, 96(2):405–410.

Campbell, J. Y. and Shiller, R. J. (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *The Review of Financial Studies*, 1(3):195–228. Publisher: Oxford Academic.

Campbell, J. Y. and Shiller, R. J. (1991). Yield Spreads and Interest Rate Movements: A Bird's Eye View. *The Review of Economic Studies*, 58(3):495–514.

Cho, J.-W., Choi, J. H., Kim, T., and Kim, W. (2016). Flight-to-quality and correlation between currency and stock returns. *Journal of Banking & Finance*, 62:191–212.

Cochrane, J. and Piazzesi, M. (2009). Decomposing the yield curve. *AFA 2010 Atlanta Meetings Paper*.

Cochrane, J. H. (1992). Explaining the Variance of Price–Dividend Ratios. *The Review of Financial Studies*, 5(2):243–280. Publisher: Oxford Academic.

Cochrane, J. H. and Piazzesi, M. (2005). Bond Risk Premia. *American Economic Review*, 95(1):138–160.

Devpura, N., Narayan, P. K., and Sharma, S. S. (2021). Bond return predictability: Evidence from 25 OECD countries. *Journal of International Financial Markets, Institutions and Money*, 75:101301.

Duffie, D. and Kan, R. (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance*, 6(4):379–406.

Fama, E. and Bliss, R. (1987). The Information in Long-Maturity Forward Rates. *The American Economic Review*, 77(4):680–692.

Fama, E. F. and French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1):3–25.

Fama, E. F. and French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1):23–49.

Faria, A. and Almeida, C. (2018). A hybrid spline-based parametric model for the yield curve. *Journal of Economic Dynamics and Control*, 86:72–94.

Gürkaynak, R., Sack, B., and Wright, J. (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304.

Morelli, J. M., Ottonello, P., and Perez, D. J. (2021). Global Banks and Systemic Debt Crises. Working Paper 28892, National Bureau of Economic Research.

Neumeyer, P. A. and Perri, F. (2005). Business cycles in emerging economies: the role of interest rates. *Journal of Monetary Economics*, 52(2):345–380.

Osmani, G. and Tabak, B. (2008). Characterizing the Brazilian term structure of interest rates. *International Journal of Monetary Economics and Finance*, 2(2):103–114.

Phillips, P. C. B. and Yu, J. (2005). Jackknifing Bond Option Prices. *The Review of Economic Studies*, 18(2):707–742.

Svenson, L. (1994). Estimating and interpreting forward interest rates: Sweden 1992-4. *IMF Working Paper*, 94/114.

Tabak, B. M. and Andrade, S. C. d. (2003). Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates. *Brazilian Review of Finance*, 1(1):19–43.

Tang, C. Y. and Chen, S. X. (2009). Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics*, 149(1):65–81.

Yamamoto, T. and Kunitomo, N. (1984). Asymptotic bias of the least squares estimator for multivariate autoregressive models. *Annals of the Institute of Statistical Mathematics*, 36(3):419–430.

A. Full Derivation of the Affine Model

I present the details of the multifactor affine model I estimate in the main text. In stating the model, I follow the same notation as in sections 2 and 3.

I start by defining a vector of factors X_t that evolve according to

$$X_t = \mu + \Phi X_{t-1} + e_t, \tag{A.1}$$

where $e_t \sim \mathcal{N}(0, \Sigma)$ follows a multivariate normal distribution and is independent over time.

The one-period payoff of a bond with maturity n is the price of the same bond in the following period, when it turns into $n-1$ period bond. Bonds with maturity zero pay one unit of currency. Given a stochastic discount factor M for payoffs denominated in currency units, the price of the zero-coupon bond is given by

$$P_t^n = E_t M_{t+1} P_{t+1}^{n-1}, \qquad P_{t+1}^0 = 1$$

or, taking logs and defining $m = \log M$,

$$p_t^n = \log E_t \exp\left\{ m_{t+1} + p_{t+1}^{n-1} \right\}, \qquad p_{t+1}^0 = 0.$$

Hence, a pricing theory amounts to picking m_{t+1}. Before proceeding, it is useful to re-state the pricing equation above in terms of excess returns. To do this, define the (log of the) nominal risk-free rate as usual: $i_t = \log E_t(M_{t+1})^{-1}$. Note this is the same as

$$i_t = -p_t^1. \tag{A.2}$$

So, the pricing equation above implies

$$0 = \log E_t \exp\left\{ m_{t+1} + i_t + rx_{t+1}^n \right\}, \qquad rx_{t+1}^1 = 0 \text{ a.s.} \tag{A.3}$$

("a.s." means almost surely, or for every realization of the shocks that happens with positive probability). I assume the following log discount factor:

$$m_{t+1} = -\delta_0 - \delta_1' X_t - (1/2)\lambda_t' \Sigma \lambda_t - \lambda_t' e_{t+1} \tag{A.4}$$

Heteroskedasticity of m allows for time-varying prices of risk, which is precisely what we need to reproduce time-varying risk premia as observed in the data. Having the innovations to the discount factor come from e_{t+1} rather than X_{t+1} is necessary to generate an affine solution. As we see next, λ_t governs the market price of risk and expected returns. It is linear of the factors:

$$\lambda_t = \lambda_0 + \lambda_1 X_t, \tag{A.5}$$

where λ_1 is square matrix.

Equation (A.2) gives the solution for the interest rate:

$$i_t = \delta_0 + \delta_1' X_t. \tag{A.6}$$

Replacing the discount (A.4) on the general solution (A.3) yields

$$E_t rx_{t+1}^n = -(1/2)\text{var}_t(rx_{t+1}^n) + \text{cov}_t(rx_{t+1}^n, e_{t+1}')\lambda_t. \tag{A.7}$$

The covariance term represents the amount of risk of each bond, or its "beta", which is why λ_t is referred to as the market price of risk.[1] Given the solution below, the conditional variance and covariance terms

[1] In the beta representation traditionally used in the empirical finance literature, an asset's beta is usually defined as the linear projection coefficient on the corresponding risk factor. Given the homoscedasticity of the innovations in the model, the coefficients emerge by simply left-multiplying λ_t by $\Sigma^{-1}\Sigma$. Then, $\text{cov}(rx^n, e')\Sigma^{-1} = \text{cov}(rx^n, e')E(ee')^{-1}$ becomes the "usual"

are both time invariant, and we can thus drop the t subscripts. All variation in the risk premium $E_t rx_{t+1}^n$ comes from λ_t.

The linear structure implies a solution for equilibrium prices that is linear in the factors. The same is true for forward rates:

$$p_t^n = A_n + B_n' X_t \tag{A.8}$$
$$f_t^n = A_n^f + B_n^{f'} X_t. \tag{A.9}$$

The relationship between forward rates and prices $f_t^n = p_t^{n-1} - p_t^n$ implies $A_n^f = A_{n-1} - A_n$ and $B_n^f = B_{n-1} - B_n$. Using our definition (A.2) of the nominal risk-free rate and the discount factor (A.4), we also find that $-A_1 = A_1^f = \delta_0$ and $-B_1 = B_1^f = \delta_1$. By replacing the definition of forward rates on the definition of excess returns $rx_{t+1}^n = p_{t+1}^{n-1} - p_t^n - i_t$, and that on the pricing equation (A.7), we find expressions that allow for the computation of the solution coefficients recursively:

$$\begin{aligned} A_n^f + B_{n-1}'\mu^* &= -(1/2)B_{n-1}'\Sigma B_{n-1} + \delta_0 \\ B_n' &= -\delta_1' + B_{n-1}'\Phi^* = -\delta_1'(I + \Phi^* + \Phi^{*2} + \cdots + \Phi^{*n-1}) \\ B_n^{f'} &= \delta_1'\Phi^{*n-1}. \end{aligned} \tag{A.10}$$

In the equations above, I define

$$\begin{aligned} \mu^* &= \mu - \Sigma\lambda_0 \\ \Phi^* &= \Phi - \Sigma\lambda_1. \end{aligned} \tag{A.11}$$

These coefficients characterize the *risk-neutral* dynamics of the factors X_t. To see this, replace the computed coefficients in (A.10) in (A.9):

$$\begin{aligned} f_t^n &= -(1/2)B_{n-1}'\Sigma B_{n-1} + \delta_0 - B_{n-1}'\mu^* + \delta_1'\Phi^{*n-1}X_t \\ &= -(1/2)B_{n-1}'\Sigma B_{n-1} + \delta_0 + \delta_1'\left[(I + \Phi^* + \cdots + \Phi^{*n-1})\mu^* + \Phi^{*n-1}X_t\right] \\ &= -(1/2)B_{n-1}'\Sigma B_{n-1} + \delta_0 + \delta_1' E_t^* X_{t+n-1} \\ &= -(1/2)B_{n-1}'\Sigma B_{n-1} + E_t^* i_{t+n-1}. \end{aligned}$$

The operator E_t^* above computes expected value using the probability measure under which

$$X_t = \mu^* + \Phi^* X_{t-1} + e_t, \qquad e_t \sim \mathcal{N}(0, \Sigma).$$

One can formally define such probability measure through the Radon-Nikodyn process $\exp(\chi_t)$, which evolves according to $\chi_{t+1} = \chi_t + m_{t+1} + i_t$. Then, for any random process Y, $E_t e^{\chi_{t+1} - \chi_t} Y_{t+1} = E_t^* Y_{t+1}$. In particular, (A.3) becomes $0 = \log E_t^* \exp rx_{t+1}^n$, which justifies the "risk-neutral" designation of this alternative probability measure, which differs from the data-generating or "true" measure defined in (A.1). The algebra above shows through a different route that, under the risk-neutral probability measure, there is no term-premium: the forward rate coincides with the expected interest (plus a Jensen's inequality term).

B. Estimation by Linear Regressions

Adrian et al. (2013) provide a method of estimating affine models using linear regressions only. I adapt their procedure to my specification, but the general method is the same.

Start by running OLS on the vector autoregression (9). That yields estimates for μ and Φ, as well as sample residuals $\hat{e}$, which we can use to estimate Σ (this is also how I estimate Σ in the baseline

beta, and $\Sigma\lambda_t$ becomes the market price of risk.

RHS Forward Rates	Adjusted R^2 on rx^n_{t+1} regression			
	$n=2$	$n=5$	$n=10$	Avg Return
All	0.50	0.52	0.57	0.54
1,2,3,4,5	0.44	0.48	0.53	0.50
2,4,6,8,10	0.37	0.46	0.52	0.49
2,4,6	**0.32**	**0.44**	**0.51**	**0.47**
2,10	0.05	0.18	0.31	0.23
2,6	0.12	0.33	0.47	0.40
6,10	0.32	0.38	0.41	0.39

The table reports the *unadjusted* R^2 for regression models $rx^n_{t+1} = \alpha + \beta' f_t + \varepsilon^n_{t+1}$. Each row contains results for a different set of right-hand side forward rate maturities.

Table 6: R^2 per model and bond maturity

	Return	Level	Slope	Curve
Return	1	0.60	0.63	0.21
Level		1	0.22	0.02
Slope			1	0.03
Curve				1
Standard Deviation (%)	2.74	4.54	1.22	0.60

I estimate (9) by OLS and compute residuals $\hat{e}_t$. The table reports its estimated correlation matrix and, in the last row, standard deviation.

Table 7: Correlation Matrix of Factor Shocks

estimation procedure).

I then run the multivariate regression

$$rx_{t+1} = \mathcal{A} + \mathcal{B}\hat{e}_{t+1} + \mathcal{C}X_t + u_t$$

where rx stacks the excess returns of all bonds. The exponential-affine model implies that the covariance of excess returns with innovations is $B\Sigma$, where B is the matrix that stacks B'_{n-1}. Therefore, $\mathcal{B}$ provides an estimate of B.

The last step is to estimate price-of-risk parameters by running cross-sectional regressions. By matching equation (14) with the linear regression above, one can see that

$$\begin{aligned}\mathcal{A}_n + (1/2)B'_{n-1}\Sigma B_{n-1} &= \lambda'_0(\Sigma B_{n-1}) + \text{error}\\ \mathcal{C}'_n &= \lambda'_1(\Sigma B_{n-1}) + \text{error}\end{aligned}$$

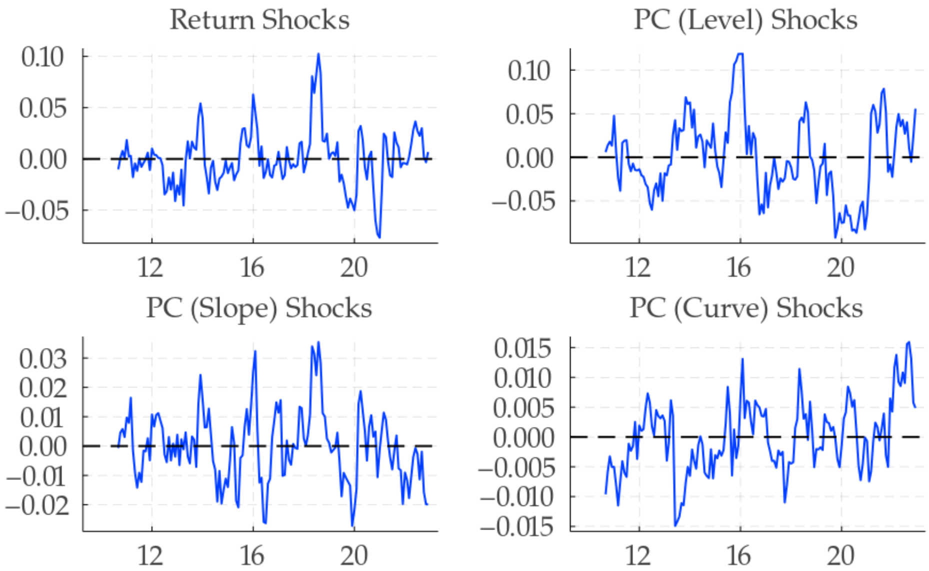
where $\mathcal{A}_n$ is the n-th entry of vector $\mathcal{A}$ and $\mathcal{C}_n$ is the n-th row of $\mathcal{C}$. In these regressions, the estimates of (ΣB_{n-1}) are the indepedent variables. Each bond maturity n is a "data point", which is why I refer to them as a cross-sectional regressions. They estimate λ_0 and λ_1 by projecting expected returns bond-by-bond (the $\mathcal{C}_n$) to estimated loadings B.

This step of the procedure is the analogous to my cross-sectional regression that selects $\lambda_1(1)$ to match (14) to the empirical excess return model (10).

Having estimated λ_0 and λ_1, we complete the estimation by calculating risk-neutral parameters μ^* and Φ^* using (15).

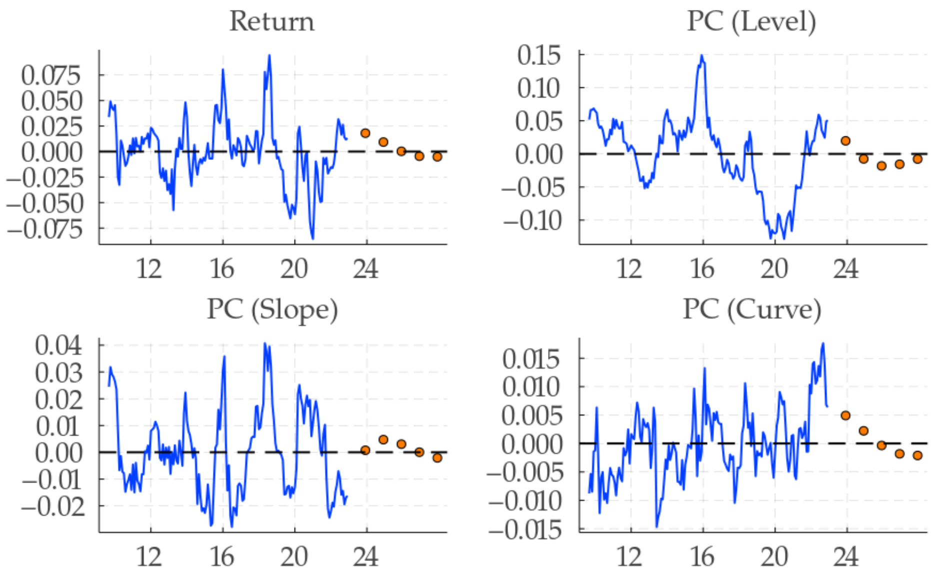
C. Other Results

This section contains additional tables and figures.



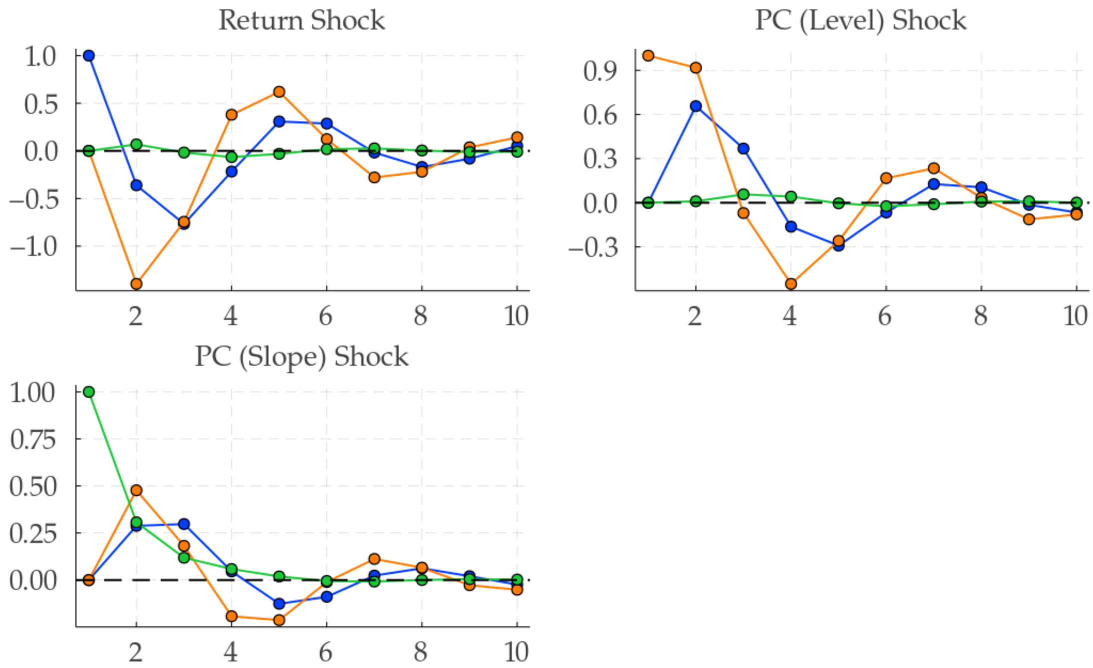
Notes: I estimate (9) by OLS. The figure plots the estimated residuals $\hat{e}_t$.

Figure 8: Estimated Factor Shocks



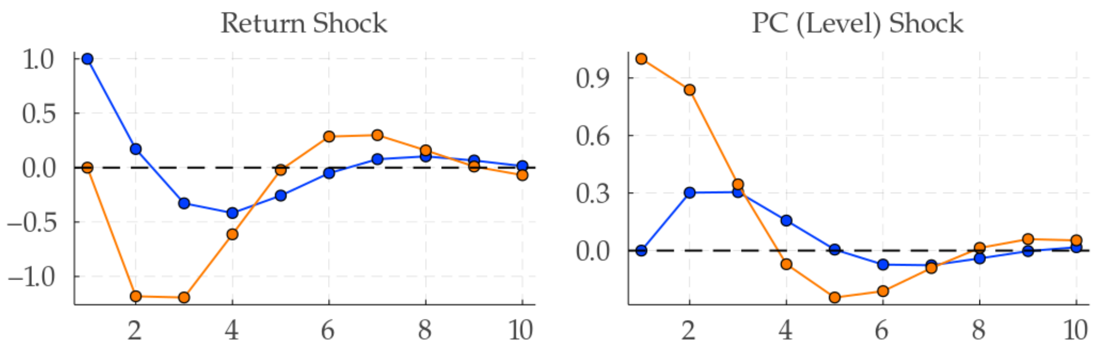
Notes: Each panel plots the time series of the models' factors (blue curves). The orange circles represent the models' one-year ahead forecast using the baseline estimation parameters.

Figure 9: Model Factors and their Forecast



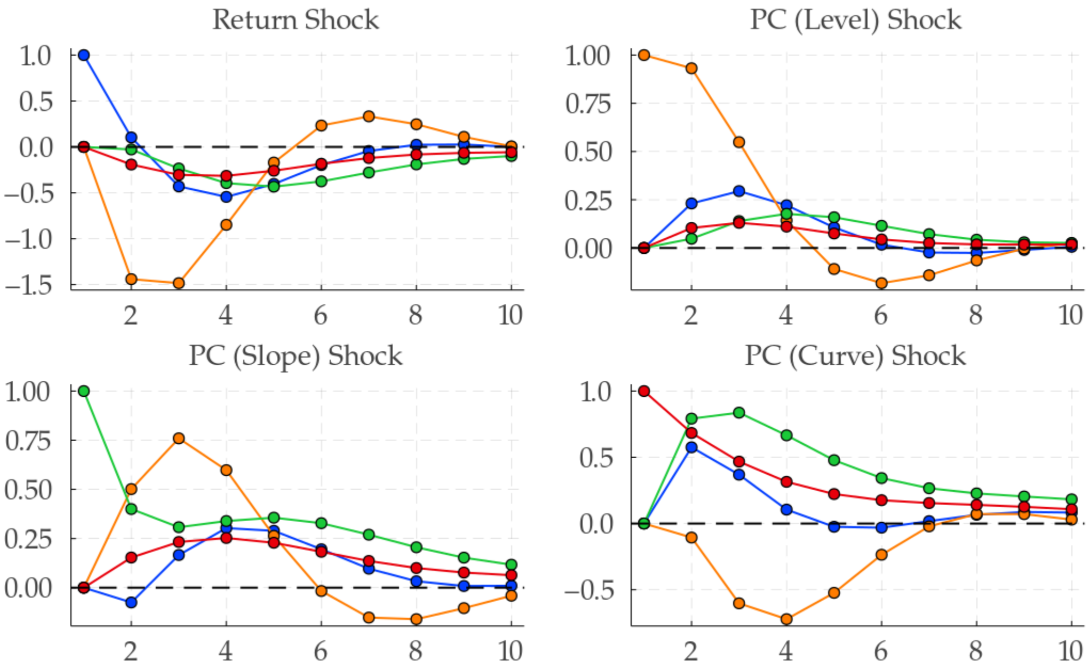
Notes: Impulse response function of model dynamics under the true probability measure. Blue = return factor. Orange = level factor. Green = slope factor. In this specification, I assume X contains three factors only.

Figure 10: Data-Generating Dynamics Φ - 3-Factor Model



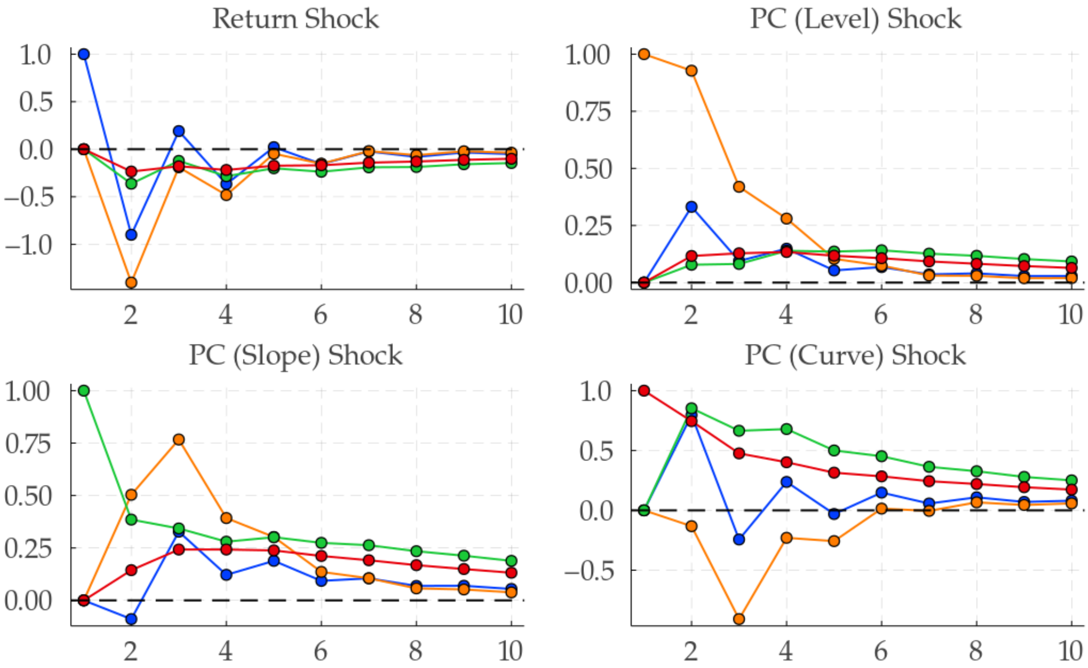
Notes: Impulse response function of model dynamics under the true probability measure. Blue = return factor. Orange = level factor. In this specification, I assume X contains two factors only.

Figure 11: Data-Generating Dynamics Φ - 2-Factor Model



Notes: Impulse response function of model dynamics under the true probability measure. Blue = return factor. Orange = level factor. In this specification, I assume the law of motion of X contains a constant.

Figure 12: Data-Generating Dynamics Φ - Model with Free μ



Notes: Impulse response function of model dynamics under the true probability measure. Blue = return factor. Orange = level factor. In this specification, I assume that only return factor shocks have a non-zero market price of risk (as opposed to level factor shocks in the baseline specification).

Figure 13: Data-Generating Dynamics Φ - Return Factor Priced