A Decade of Term Structure in Brazil: Return Predictability and Yield Decomposition

Livio Maya*

Banco Safra

October 21, 2021

Abstract

Long-duration public bonds denominated in local currency have been traded with liquidity in Brazil since the late 2000s. I use the ten years or so of (now) available data to study return predictability and bond risk premia following the steps already taken for developed markets. I compute a return factor and use it along with term structure principal components to estimate a multifactor affine model that allows for the decomposition of forward rates. The following results emerge: (1) Fama-Bliss regressions support the expectations hypothesis in Brazil; (2) Despite (1), bond returns are predictable; (3) Like for the US, a single combination of forward rates predicts annual returns with high precision (R^2 reaches 48%); (4) Level shocks to the term structure have, on average, *negative* market price of risk - public bonds are, thus, inherently *risky* investments, and marginal utility *rises* with interest rate; (5) The discount factor places heavy weight on sequences of shocks that lead to prolonged periods of high marginal utility - investors fear that times of low consumption such as the 2015-2016 recession will last longer than they do on average; (6) Shocks are long-lived, and expected long-term interest rates have fluctuated almost as much as forward rates; and (7) There is a non-trivial *term structure of risk premia*: low returns today are consistent with large returns foreseen in the future, and the term premium is therefore heterogeneous along the term structure.

1 Introduction

Following the end of hyperinflation in the mid-1990s, successive Brazilian presidential administrations took steps in the direction of extending the average maturity of public debt and changing its composition to encompass primarily fixed-income, local currency bonds. This change in the profile

*I deeply thank my former professors at Stanford University, Prof. Monika Piazzesi and Prof. John Cochrane, whose brilliant research has heavily inspired the work I document in this article. I also thank Banco Safra for supporting me, as well as Joaquim Levy and Eduardo Yuki for the numerous discussions, and Matheus Wirth for the excellent research assistantship.

of public debt gave rise to a previously non-existent liquid secondary market of nominal bonds and, by extension, a measurable term structure of implied interest rates.

Availability of such information - in particular long-term yields - is of critical importance for academic finance research. The prevailing message of modern literature is that returns are predictable, and compensate investors for holding certain risks. Return predictability nevertheless is only there if we measure returns over long periods of time - years, usually. The task of the finance studier is to measure expected returns (the risk premium), how it changes over time, and explain which risks are being rewarded. Until a few years ago, none of these questions could be answered in the case of Brazil. The necessary data (or market. or bond.) just wasn't there. Now it is[1].

By the time I write this paper, bonds with duration of up to ten years have been traded for about ten years. The intent of this article is to use the now available data to follow the research steps taken in the cases of countries with older markets, and study predictability and bond risk premia in the still young Brazilian market. I am interested in a series of questions: are returns predictable? Does the expectations hypothesis hold? How does risk premium change over time? What clues do current prices give about it? Can we decompose the yield curve and measure the term premium? How does it relate to movements in the term structure? Does it even make sense to refer to a single "term premium", or are term premia heterogeneous in the maturity dimension?

The paper has three parts. In the first part, I describe and analyze the data. In the second part, I study return predictability. In the third part, I estimate an exponential-affine model that ties down my empirical findings and allows for yield decomposition.

I use data provided by the Brazilian Financial and Capital Markets Association (*ANBIMA*) for the period 2010-2021. It collects market prices from publicly traded government bonds and provides estimated parameters from a fitted Svenson (1994) model, which I, lacking better alternatives, treat as data. I document that some of the better know facts about the term structure of interest in other economies hold in Brazil as well. For example, mean forward rates (or yields) increase with maturity, while their volatilities decline. I also use the eigenvalue decomposition to compute principal components. Like in the US market (Litterman and Scheinkman (1991)), three principal components (the *level*, *slope* and *curve* factors) capture almost all variation in yields, a fact I explore when choosing model factors.

I also document stylized facts about bonds' excess returns and - moving now to the second part of the paper - look for variables that can forecast them. Throughout the paper, I focus only on annual returns, given that, at higher frequencies, returns are hardly predictable. To test the expectations hypothesis - a natural starting point -, I start by running Fama and Bliss (1987) style regressions. In the US case, estimates decisively *reject* the expectations hypothesis, especially in the short run. Forward spreads - the difference between forward rates and the short-term interest - do not predict higher interest in the future. They predict higher returns to investors. In the Brazilian case, I find the exact opposite. In the 2010-2021 period I analyze, forward spreads were associated

[1]For the reasons highlighted in this paragraph, papers focusing on the Brazilian term structure are rare. For applications using older datasets of bonds with short maturities, see Tabak and Andrade (2003) and Osmani and Tabak (2008).

with interest growth, not with excess returns. I attribute that to highly volatile interest rates in Brazil, a variable the literature claims to be "sticky" in the US.

The results of the Fama-Bliss regressions, however, do not prove returns are unpredictable. On the contrary, I run linear regressions and find that forward rates have a high capacity of forecasting them. I set as right-hand variables the rates of maturities two, six and ten years (increasing the number of maturities leads to multicollinearity issues), although other combinations also work. R^2s are large, ranging from 22.5% for the two-year bond excess return to 48.6% for the ten-year bond return. In addition, I find that, just like Cochrane and Piazzesi (2005) report in their seminal study of the US market, returns' loadings on forward rates appear to be "scaled" versions of each other across regressions. Based on this finding, I compute a *single return-forecasting factor* for the Brazilian nominal bond market. The factor predicts returns with almost the same precision as the regressions and can be stated (up to sum and scale) in an intuitive, one-parameter formula:

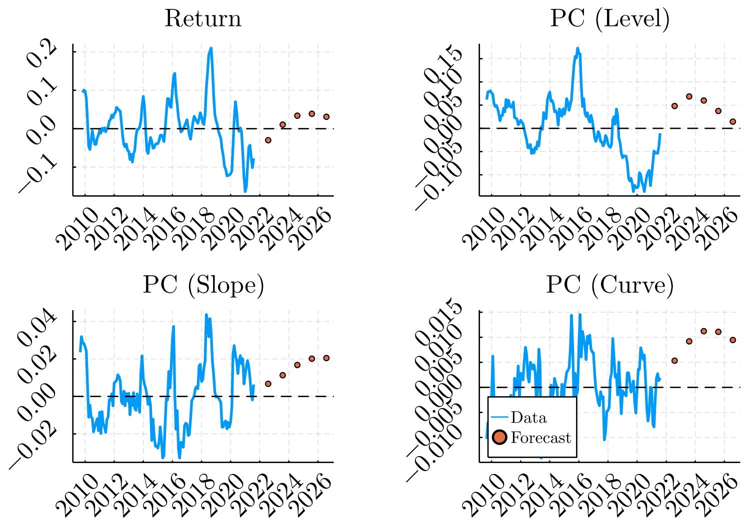
$$\text{Return Factor} = \text{Short-Term Rate} + 2\,(\text{Short Slope} - \text{Long Slope})\,.$$

There is nothing special with the forward rates I select (maturities 2, 6 and 10). Other proxies for the short-term rate, short slope and long slope also work.

I study each component of the return factor and describe how they were important to predict returns in the 2010s. For instance, the short-term rate is critical to describe the large returns following the 2016 recession, while the short slope captures the large returns in 2019-2020. Our return factor is not spanned by the three main principal components. In addition, I show that, given the return factor, traditional level, slope and curve factors offer no additional return forecasting power.

In the third part of the paper, I reverse-engineer a stochastic discount factor using the now well-known exponential-affine model framework (Duffie and Kan (1996)) to explain the term structure of interest rates. The model allows the estimation of the return factors' law of motion along with that of the other factors. As I do not offer an economic theory for the discount factor, I choose as factors, in addition to the return factor, the level, slope and curve series formed from the principal components decomposition of the term structure. For future reference throughout the paper, figure 1 plots the time series of each of the four factors. The red dots indicate my forecast according to the estimated model and conditional on the information contained in the sample.

To estimate the model, I follow Cochrane and Piazzesi (2009). I impose restrictions on the market price of risk, which reduce our ability to explain prices but allow for a cleaner connection between the data-generating (or "actual") probability measure and the risk-neutral probability measure. The first restriction is that *the market price of risk varies only according to the return factor*. Because the model is conditionally homoskedastic, this means that the remaining factors do not forecast returns given the return factor. Just like regression estimates suggest. The second constraint is that the only priced risk is the shock to the *level factor*. This restriction follows from the comparison between the schedule of conditional return/shocks covariances and expected returns. Both return and level factors' shocks match the term structure of expected returns. I impose the



Notes: Each panel contains one of the four factors of the model. The first factor is built from a combination of forward rates to maximize return forecastability, hence its name. The other three factors are the main principal components calculated from the first ten years of forward rates. They are usually interpreted as level, slope and curve factors, following their expected effect on forward rates of different maturities. The orange markers indicate the path projected for each factor using the model's baseline estimation.

Figure 1: Model Factors and their Forecast

constraint on the latter due to model fitting. In all, my constraints are similar to Cochrane and Piazzesi (2009): the return factor drives the risk premium; the risk premium compensates investors for exposure to level factor shocks. These constraints reduce twenty parameters governing the market price of risk to two.

I estimate the model in three steps. First, I minimize pricing errors by selecting the *risk-neutral* dynamics of the factors. Second, I the set market price of risk loading on the return factor so that the model approximates the cross-sectional linear regressions I estimate. Third, I use the risk-neutral dynamics and the market price of risk to infer factors' law of motion under the "true" or data-generating dynamics.

Under the risk-neutral probability measure, factors must be persistent, since the level factor explains about 90% of variation in forward rates. I find that model factors are persistent under the data-generating measure as well. This leads to model-implied volatile *expected future interest rates*. News about today's economy significantly affect agents' perception of interest in the mid/long term, a result we do not find *via* OLS estimation of factors' law of motion (as in Adrian, Crump, and Moench (2013)). Empirically, I calculate that long-term expected interest rates have fluctuated during the 2010s almost as much as their corresponding forward rates.

Our estimates also provide macroeconomic interpretation to the factors. For example, I estimate the market price of risk of level shocks (the only ones assumed to be priced) to be negative, on average, a finding similar to Joslin, Priebsch, and Singleton (2014). Hence, investors that purchase payoffs that are positively correlated with them pay a premium, and expect lower returns. Such payoffs thus act like *insurance*, which indicates that *positive level shocks or a rising term structure* *in Brazil correspond to a rising marginal utility*, at least on average. Given that bonds tend to have a worse performance in times of rising yields, they are therefore *risky* investments in a theoretically rigorous use of the word.

Comparison between factors' dynamics under risk-neutral and true probability measures is also useful to characterize the stochastic discount factor, and thus marginal utility, directly. Under the actual probability measure, an increase in the level factor forecasts an increase in the return factor. An increase in the return factor, in turn, *reduces* the level factor, and renders stable dynamics. Under the risk-neutral measure, however, this second step fails. An increasing return factor does *not* lead to lower yield levels. Hence, shocks to yields (marginal utility) last for a long time. By construction of risk-neutral probability measures, such difference in factors' dynamics put in evidence the realization of states of nature that feature comparatively higher marginal utility. I therefore conclude that investors' utility is particularly sensible to sequences of shocks that lead to prolonged recessions, like 2015-16. Such conclusion is in accordance with existing literature emphasizing the importance of long-lasting shocks to the business cycle of emerging markets (see Aguiar and Gopinath (2006) and Aguiar and Gopinath (2007)).

Finally, having characterized factors' dynamics and the market price of risk, I decompose the term premium. The term premium is determined by the evolution of risk premium over time. I show that the *term structure of risk premium* is non-trivial (using Cochrane and Piazzesi (2009) expression) in the sense that different shocks have different effects on risk premium in the short run and in the long run. Living in a period of low expecetd returns does not mean agents cannot foresee high returns in the horizon. If they do, the term premia of long-term bonds will be large, while that of short-term bonds will be low. Hence, we cannot refer to a single term premium. The term premium varies with, well, term.

I take this paper to be an initial step in the direction of reproducing for Brazil a lifetime of research already established for developed markets, especially for the US. To the best of my knowledge, the studies of predictability and the fitting of the exponential-affine model I propose in this article have not yet been applied to Brazil - possibly due to lack of proper data. Given its young market, sampling uncertainty is obviously a big concern, and a concern I can only partially address by choosing the most parsimonious specifications and not sticking to results that are more sensitive to small sample issues.

2 A Decade of Term Structure in Brazil

2.1 Notation

As I work with annual returns, one unit of time in my notation represents one year.

Let p_t^n denote the log of the period t price of a zero-coupon nominal bond that comes due n

periods into the future. Let f_t^n be the log of the forward rate n periods ahead[2]:

$$f_t^n = p_t^n - p_t^{n-1}, \tag{1}$$

so that $i_t = f_t^1 = -p_t^1$ is the risk-free rate (I use one-year forward rate, risk-free rate and interest rate interchangeably). Let rx_{t+1}^n be the log of the return of a n-period bond in period $t+1$ (when it becomes of duration $n-1$) in excess of the risk free rate $i_t = f_t^1 = -p_t^1$, or simply the excess return:

$$rx_{t+1}^n = p_{t+1}^{n-1} - p_t^n - i_t. \tag{2}$$

Data on the prices of zero-coupon bonds is sufficient to build a complete term structure of forward rates and the excess return on bond holdings.

2.2 Data and Macroeconomic Context

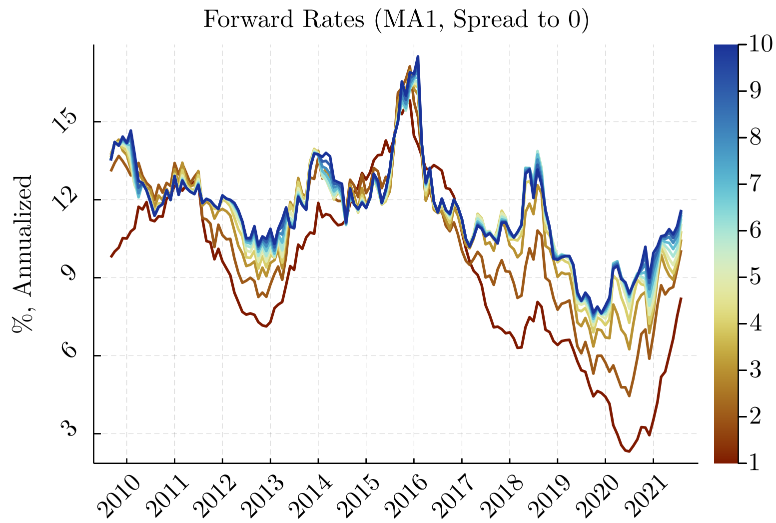
The Brazilian Financial and Capital Markets Association (*ANBIMA*) collects market price data of nominal bonds issued by the Brazilian federal government. Like other countries, at any given point in time outstanding Brazilian bonds do not cover a full range of maturities, and not all bonds are zero-coupon. 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. Parameters are re-calculated every day to approximate the discount applied to the more liquid bonds outstanding[3] in the Brazilian secondary market. I use them to build monthly observations of the prices of bonds with maturities $n = 1, 2, \cdots, 10$ years. The closest analogue to this estimation approach for the US is studied by Gürkaynak, Sack, and Wright (2007). Their estimated term structure of interest is, for the most part, similar to the more commonly used Fama-Bliss dataset. I treat ANBIMA's estimates as data throughout the paper, and highlight when the fact that I am using a fitted function can be a caveat to the analysis.

Out term structure "data" covers the period Sep-2009 to Aug-2021, and each data point corresponds the last business day of the corresponding month. I apply logs and use (1) and (2) to build the time series of forward rates and excess returns. Note that returns are over a *one-year* holding of the bond. Figures 2 and 3 present the corresponding plots. Each curve represents a different maturity (in years), with the color bar on the right indicating which is which.

Unfortunately, the time range of the data does not cover multiple swings of the business cycle, and thus some description of the economic context can be useful. The Brazilian economy experiences a decade of strong economic growth in the 2000s, and quickly recovers from the Great Recession. This is when the dataset begins. Starting in 2011, growth begins to wane, and monetary policy rates - best but not perfectly described in figure 2 by the one-year forward - decline. A strong recession hits the economy in 2015 and the business cycle reaches a trough in the following year.

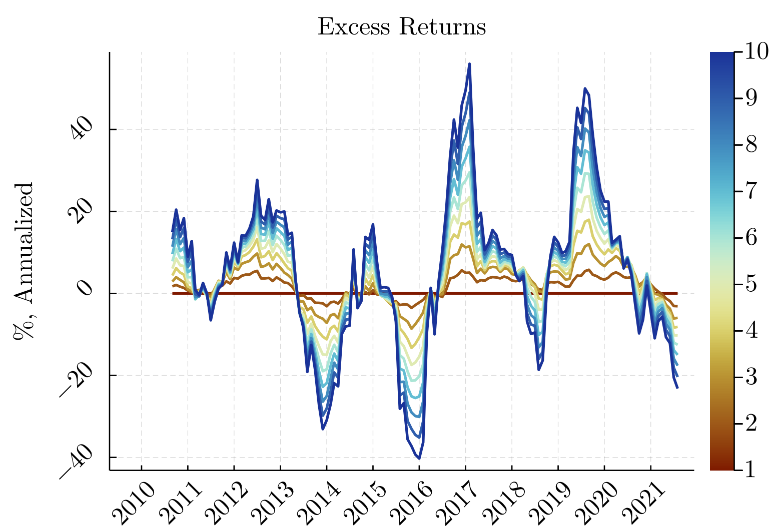
[2]I could just as well focus on yield rates instead of forward rates. I opt for the latter since it has a more direct connection to expected interest rates in the model, which facilitates exposure.

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



Notes: I build the term structure using ANBIMA's estimates of the parameters from the Svenson model for yields. The estimates attempt to minimize the distance between theoretical and observed yields from an array of fixed-income bonds. I calculate forward rates in logs. Monthly data. Each data point corresponds to the last observed rate in the corresponding month. Each curve corresponds to a different maturity. The color bar indicates the maturity, in years.

Figure 2: Data - Forward Rates



Notes: Data from ANBIMA (see notes to figure 2). Monthly observations of one year returns. Each curve corresponds to a different maturity. The color bar indicates the maturity, in years.

Figure 3: Data - Excess Returns

The Brazilian central bank reacts to capital flights and soaring inflation by raising policy rates. Economic growth, mild ever since, resumes in 2017. Along with fiscal consolidation efforts by the public sector, the central bank slowly reduces policy rates to unprecedent levels. The one-year forward rate reaches its minimum by mid-2020, at the early stages of COVID-19 Pandemic (central bank target rates only begin to increase in 2021 - they do not always coincide).

2.3 Unconditional Moments

Figures 2 and 3 gives us a first impression of the dynamics of yields and returns. However convenient, visual evaluation alone can mask important features of the data.

Tables 1 and 2 present some sample moments of forward rates and excess returns. I also separately report moments for forward *spreads*, defined as differences between forward rates of different maturities. It is of course possible to infer their moments from the moments of the underlying forward rates, but given their importance when building the return factor, I include them here as well. In the tables, spread are depicted by the letter "S": S(10-1) = f(10)-f(1) and so on. Table 1 focuses on each series individual moments (mean, volatility and persistence); table 2 focuses on cross-section correlations. I report all values in percentage units.

Starting with table 1, the term structure of average forward rates is positively sloped: long rates are, on average, larger than short ones. These averages range between about 9 and 12% per year. Short rates, on the other hand, are more volatile. Our estimated standard deviations range between 3.3% for the risk-free bond and 1.8% per year. All forward rates have a positive one-year autocorrelation coefficient, but short-duration bonds are more persistent, with coefficients above 0.5 in the cases of the one and two-year maturities. These autocorrelations uniformly decline to the 20-25% range when I look at a two-year horizon and, at a three-year horizon (not reported), they all turn negative.

Note that in the short end of the curve, moments are rapidly changing. Between maturities one and four years, the average rate increases from 9.2 to 11.3%, the standard deviation decreases from 3.3 to 2% and the one-year autocorrelation falls from 0.65 to 0.35. But in the long end of the term structure, moments become considerably more homogeneous across bonds of different maturities. Between maturities four and ten years, the mean rate goes from 11.3 to 11.6%, standard deviations from 2 to 1.9% and autocorrelation coefficients from 0.35 to 0.27.

Moving on to forward spreads, the average 10y-1y spread is 2.4%, with the bulk of this difference coming from the spread in the first five years of the term structure, as one might expect from the remarks made in the previous paragraph. Additionally, the short spread S(5-1) shows an estimated standard deviation between four and five times greater than that calculated for the long spread S(10-5). This observation will be useful in the process of building a return-forecasting factor. Finally, while S(5-1) presents some persistence, with one and two-year autocorrelation coefficients of 0.36 and 0.16, respectively, S(10-5) appears to quickly revert to its unconditional mean.

Looking now at excess returns - the table ignores R(1) since it equals zero by construction - we see that the "short-term heterogeneity/long-term homogeneity" trait of forward rate moments is,

	Mean %	**St Dev** %	**AC(1y)** %	**AC(2y)** %
f(1)	9.2	3.3	65.4	20.1
f(2)	10.4	2.7	54.8	22.2
f(3)	11.1	2.2	42.5	24.7
f(4)	11.3	2.0	34.6	23.2
f(5)	11.4	1.9	30.5	21.6
f(6)	11.4	1.9	28.7	20.8
f(7)	11.5	1.8	27.9	20.6
f(8)	11.5	1.8	27.5	20.6
f(9)	11.6	1.8	27.1	20.6
f(10)	11.6	1.9	26.7	20.5
S(10-1)	2.4	2.2	34.3	16.4
S(5-1)	2.1	2.1	35.8	16.3
S(10-5)	0.2	0.5	8.7	-11.3
Rx(2)	1.6	2.5	4.5	-2.1
Rx(3)	2.6	4.9	-8.5	3.7
Rx(4)	3.3	7.2	-18.7	8.3
Rx(5)	3.7	9.3	-25.9	10.9
Rx(6)	4.1	11.4	-31.1	12.5
Rx(7)	4.4	13.5	-34.8	13.5
Rx(8)	4.8	15.6	-37.6	14.2
Rx(9)	5.1	17.7	-39.7	14.8
Rx(10)	5.4	19.9	-41.4	15.3

Notes: 2010-2021 Brazilian term structure data. All values in percentage units. "f" means forward rate, "S(a-b)" means spreads (f(a)-b(b)), and "Rx" means excess return. AC(1y) means one-year Pearson autocorrelation, same for AC(2y).

Table 1: Table of Sample Moments

to a large extent, gone. The average excess return increases from the short to the long end of the term structure along with its volatility. Assuming my sample is sufficiently large to approximate populational moments well, estimated means reflect positive unconditional risk premia: holding public bonds provide investors with a return greater than the risk-free rate, on average. This suggests that bonds' payoffs are *risky* in the sense that they covary negatively with marginal utility, a claim we cannot put to the test without a proxy for marginal utility - this is where we are going. Estimated unconditional risk premia range from 1.6% per year for two-year bonds to 5.4% per year for ten-year bonds. In a far looser use of the word *risk*, what we can say from the data is that long-duration bonds are riskier because their return is more volatile. Estimated standard deviations of excess returns increase almost linearly from the two to the ten-year duration bonds, with increments from term to term always between 2 and 2.5%.

Finally, table 1 shows that, while returns from short-duration bonds display almost no persistence, returns from long-duration bonds display *negative* persistence. One-year autocorrelation coefficients become increasingly negative, and reach an economically significant -0.41 value for the ten-year bond. In the case of these bonds with long duration, a good year for investors is usually followed by a bad year. In addition, for the two-year autocorrelation, coefficients move back into positive territory, despite having lower absolute values, which suggests that excess return dynamics displays a "wave" shape at an annual frequency. Good years are followed by bad years, and then good years again.

Moving to table 2, we focus now on correlations. Each entry of the table contains the Pearson correlation between the row variable and the column variable. Including all maturities would take a lot of space and add little insight, so I select a subset of them that hopefully represents well the short, medium and long end of the curve. An important clarification: I compare forward rates measured in period t with returns realized in period $t+1$ (one year later), *not* t. As we saw on table 1, excess returns are negatively correlated over time. So, the comparison between time-t rates and time-t returns yields different results. I stick to the former comparison since our interest is how forward rates relate to the performance of portfolios formed by the time they are observed (the payoffs of which we observe a period later).

Starting with forwards, we see that rates are highly correlated across maturities, as expected. Two additional comments: 1. the closer the maturities, the more correlated are the corresponding rates; 2. the one-year forward - our measure of a risk-free rate - is slightly *less* correlated with the rest of the curve than its counterparts, although coefficient estimates remain elevated. Similarly, excess returns are highly correlated across maturities, with estimated coefficients declining in the duration difference.

I also observe that forward rates and excess returns are positively correlated: when prices are low compared to their unconditional average, returns tend to be high. The correlation strengthens for long maturity forwards and returns. In fact, forward rate correlations with the excess return of the two-year bond do not go above 0.20 for any maturity.

Finally, we look at forward rate spreads. The 10y-1y spread f(10)-f(1) is negatively correlated

	f(1)	f(5)	f(10)	Rx(2)	Rx(5)	Rx(10)	S(10-1)	S(5-1)	S(10-5)
	%	%	%	%	%	%	%	%	%
f(1)	100.0	82.1	80.0	11.7	19.1	24.0	-85.6	-85.0	-18.2
f(2)	93.9	92.7	91.8	11.4	25.8	33.5	-65.9	-65.4	-14.5
f(3)	87.8	98.0	95.7	18.0	36.4	44.8	-53.2	-50.7	-21.2
f(4)	83.9	99.7	96.6	21.4	42.2	51.4	-46.4	-42.9	-24.7
f(5)	82.1	100.0	97.2	21.7	43.9	53.9	-43.0	-39.6	-23.5
f(6)	81.4	99.9	98.0	20.6	43.4	54.1	-41.3	-38.7	-19.5
f(7)	81.1	99.5	98.9	19.0	42.1	53.3	-40.2	-38.6	-14.4
f(8)	80.9	98.9	99.5	17.4	40.5	52.1	-39.2	-38.8	-9.2
f(9)	80.5	98.1	99.9	15.8	39.0	50.9	-38.3	-38.9	-4.3
f(10)	80.0	97.2	100.0	14.5	37.6	49.7	-37.4	-38.9	0.1
S(10-1)	-85.6	-43.0	-37.4	-4.8	5.3	8.8	100.0	97.8	28.2
S(5-1)	-85.0	-39.6	-38.9	1.9	11.0	12.6	97.8	100.0	7.6
S(10-5)	-18.2	-23.5	0.1	-34.6	-29.7	-19.3	28.2	7.6	100.0
Rx(2)	11.7	21.7	14.5	100.0	91.4	81.6	-4.8	1.9	-34.6
Rx(3)	16.4	33.1	26.1	98.0	97.4	90.4	-1.3	5.1	-33.4
Rx(4)	18.1	39.7	33.0	94.6	99.5	95.0	2.6	8.7	-31.7
Rx(5)	19.1	43.9	37.6	91.4	100.0	97.4	5.3	11.0	-29.7
Rx(6)	20.0	46.8	41.0	88.7	99.7	98.7	7.0	12.3	-27.5
Rx(7)	21.0	49.1	43.7	86.4	99.2	99.4	7.9	12.8	-25.3
Rx(8)	22.0	51.0	46.0	84.5	98.6	99.8	8.5	12.9	-23.2
Rx(9)	23.1	52.5	48.0	82.9	98.0	100.0	8.7	12.8	-21.2
Rx(10)	24.0	53.9	49.7	81.6	97.4	100.0	8.8	12.6	-19.3

Notes: 2010-2021 Brazilian term structure data. All values in percentage units. "f" means forward rate, "S(a-b)" means spreads (f(a)-b(b)), and "Rx" means excess return. Each cell contains the Pearson correlation between the variables indicated in the corresponding row and column. I compare forward rate f_t^n with one-year ahead excess returns rx_{t+1}^n.

Table 2: Table of Sample Moments - Correlations

with the rates of *all* ten maturities. The term structure of interest rates tends to have a higher slope when yields are low, a common pattern which we can attribute to the fact that long and short duration rates have a strong, positive correlation and that long rates are less volatile:

$$\mathrm{cor}(f^{10}, f^{10} - f^{1}) \propto \sigma(f^{10}) - \mathrm{cor}(f^{10}, f^{1})\sigma(f^{1})$$

($\sigma(x)$ means standard deviation of x, $\propto$ means "proportional to"). The table also shows that the bulk of such correlation comes from the dynamics of the short spread, measured in the tables by f(5)-f(1). The long spread, besides displaying reduced volatility and autocorrelation when compared to the other half of the slope, is also less correlated (in absolute terms) with forwards rates. In fact, by looking at the cross-correlation between our spread measures (right column, middle rows), we verify that the dynamics of the "full" slope S(10-1) is very similar to the dynamics of its first half S(5-1). Both display similar volatilities (2.2 and 2.1%) and a correlation coefficient that borders 0.98. On the other hand, I estimate the correlation of either of them with the long spread S(10-5) to be far lower: 0.08 for the short spread, 0.29 for the full slope. Lastly, the short spread (and so, almost as a consequence, the full slope) is positively correlated with excess returns of short-duration bonds. When prices are low compared to current stances of short-term rates, returns tend to be high. Like with forward rates, correlation coefficients increase as we look at the return of bonds with longer maturities. Nevertheless, note that such correlations are somewhat weaker (never above 0.13), especially if we compare them with the coefficients calculated for forward rates of maturities five and ten (often above 0.40).

As for the long spread S(10-5), its relationship with excess returns is not in this case a mere toned down version of the returns/mid-short spread correlation structure. All coefficients are larger in absolute value, negative and decrease (in absolute value) for longer maturities. This feature of the data provides a first indication that different parts of the term structure slope convey different information regarding expected returns. This is also consistent with the "tent"-shaped pattern of OLS coefficients found by Cochrane and Piazzesi (2005) in their return regressions on forward rates. The first half of the tent says that a larger short slope predicts higher returns; the second half says that a larger short slope predicts lower returns. The analysis of unconditional moments for the Brazilian term structure suggests that a similar pattern applies, with the difference being that I look at maturities of up to ten years. Such similarity will show up again when we build the return factor.

2.4 Principal Components

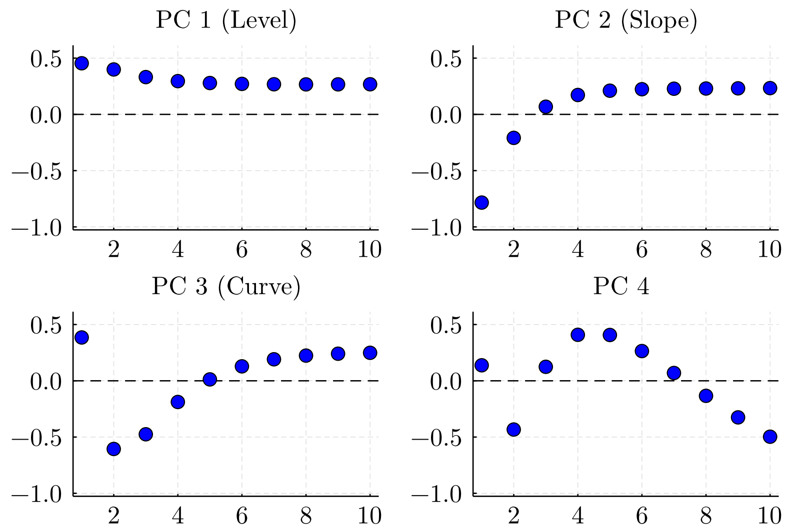
We can build principal component factors that describe the movements of forward rates of different maturities using an eigenvalue decomposition of their covariance matrix (Litterman and Scheinkman (1991)). A well-known result that holds for the US bond market is that the first three principal components explain most (99%+) price variation.

The same is true for the Brazilian term structure, as one can guess from the strong correlation between rates of different maturities reported in table 2. Table 3 reports the cumulative share of

	Cum Var	Std Dev
	%	Annual %
PC 1 (Level)	92.9	6.7
PC 2 (Slope)	98.9	1.7
PC 3 (Curve)	99.5	0.6
PC 4	100.0	0.5
PC 5	100.0	0.1

Notes: I calculate the principal components of the forward rates' data using the eigenvalue decomposition of the sample covariance matrix. The first column reports the cumulative variance of each component. The second column reports their standard deviation. All values reported in percentage units.

Table 3: Principal Components - Cumulative Variance Share



Notes: Horizontal axis indicates the forward maturity. I calculate the principal components of the forward rates' data using the eigenvalue decomposition of the sample covariance matrix. Each plot contains the weights $w^{(i)}$ of each forward rate in the calculation of the indicated principal component, for example $pc_t^{(1)} = w_1^{(1)} f_t^1 + \cdots + w_{10}^{(1)} f_t^{10}$. The weights coincide with the loadings of the forward rate on the corresponding principal component, hence the names in parenthesis.

Figure 4: Principal Component Loadings

variance for each additional principal component I consider. Three components get us over 99% of the total variance in the term structure.

Figure 4 plots the associated eigenvectors - the weights of each rate when we construct our factors. For example, the top left graph indicates that, to build the first principal component, we sum the series of forward rates of all maturities, each weighted by a positive constant ranging from 0.4 to 0.5. The principal components decomposition uses an orthonormal matrix to convert the data. For this reason, the eigenvectors I plot also correspond to the loadings of each forward rate on the corresponding factor. Focusing again on the top left graph, this means that a 1% increase in the first factor leads to a increase in each forward rate of 0.4 to 0.5%. Because yields in this case change almost uniformly, we call that the *level* factor. Analogously, the next two factors are called *slope* (or steepness) and *curve* factors. I plot the demeaned time series for each of them in figure 1.

The principal components decomposition is useful to justify the adoption of factor models when pricing the term structure. They also provide natural candidates for factors when theory does not pin them down. Indeed, to the extent that prices summarize the relevant information available to investors and how they react to it, one could assess economic news by measuring their impact on different yields. The downside, of course, is that we lose any "structural" interpretation of new information. All that matters is how the term structure of interest changes, and only in the dimensions we pay attention to (three, if we take three principal components). A new round of monetary tightening by the central bank and news of the latest primary deficit by the government are the exact same thing if both increase rates uniformly. We see the level factor jump and that's it. Similarly, two policy rate cuts of similar magnitude by the central bank account for different shocks if only one of them affects long-term yields.

One can question if it is practical to refer to economic news and events in terms of how they affect "levels" or "slopes" of the term structure. While having structural interpretation to factor movements is certainly nice, perhaps a more important question is whether it is appropriate to brand with the same name ("monetary" or anything other) two shocks that affect prices differently.

3 Return Predictability

3.1 Forward Decomposition and the Expectations Hypothesis

The prevailing paradigm of modern finance is that returns are predictable over long horizons and compensate investors for exposure to certain risks. In the context of bond markets, this idea manifests as a refusal of the *expectations hypothesis*, the claim that forward rates reflect only the expected value of future interest. Indeed, the following relationship between forward rates and expected interest and returns holds:

$$\begin{aligned} f_t^n &= E_t i_{t+n-1} + E_t r_{t,t+n-1}^n - E_t r_{t,t+n-1}^{n-1} \\ &= E_t i_{t+n-1} + E_t(rx_{t+1}^n - rx_{t+1}^{n-1}) + E_t(rx_{t+2}^{n-1} - rx_{t+2}^{n-2}) + \cdots + E_t(rx_{t+n-1}^2), \end{aligned} \tag{3}$$

where $r_{t,t+k}^n = p_{t+k}^{n-k} - p_t^n$ denotes the total return over a k-period holding of the n-maturity bond. Decomposition (3) holds *ex-post* (Cochrane and Piazzesi (2009) provide pretty art illustrating it), and so they hold *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* that groups differences in risk-premium of consecutive maturity bonds over time. The difference in price that the forward rate represents (equation (1)) needs to reflect the difference in expected payoffs for holding each bond over time until maturity. The second line of expression (3) simply groups these payoff differences according to when they occur. Note, therefore, that (3) contains *two* decompositions: the decomposition of forward rates into expected interest and term premium (on which we focus now), and the decomposition of the term premium into different risk-premum components.

If the expectations hypothesis holds, $f_t^n = E_t i_{t+n-1}$. Applying that to (3) evaluated at $n = 2$

yields $rx_t^2 = 0$. Iterating the same argument for longer maturities then yields $E_t rx_{t+1}^n = 0$ for any n under the expectations hypothesis. Returns are unpredictable.

On the other hand, if returns *are* predictable, one can assign a non-zero value to the term premium in (3), and forward rates no longer reflect only expected future interest. More than that, the expression highlights that the *time location* (or *term structure*) of risk premium matters. If we expect risk premium to increase three years from now, but not before, then the expected forward-spot spread $f_t^n - E_t i_{t+n-1}$ will be higher for maturities of four years and longer than for the two and three year maturities.

Our estimates of mean excess returns in table 1 provide an initial indication against the expectations hypothesis. On average, excess returns are large, and their size is economically significant (on average, 1.6% for $n = 2$ against 9% interest yields). Additionally, they increase by a large amount with bond duration, reaching 5.4% per year.

In the remainder of this section, I continue to study return predictability of Brazilian bonds, with a focus on predictability *conditional* on additional information that might indicate times when returns are higher or lower. I start by further investigating the expectations hypothesis through Fama and Bliss (1987) regressions of future returns and interest on current forward rates. I then proceed to search for return factors by regressing future returns on past returns, forward rates and principal components.

3.2 Fama-Bliss and Interest Variability

The forward rate is the rate at which one can lock today a one-period borrowing or loan contract in a future period. The definition suggests a connection between what we observe forward rates to be today and what we expect interest rates to be in the future[4]. In their classic study of such connection, Fama and Bliss (1987) estimate the coefficients of linear regressions of interest rate growth $i_{t+n-1} - i_t$ and realized excess returns rx_{t+1}^n on forward spreads $f_t^n - i_t$:

$$\begin{aligned} i_{t+n-1} - i_t &= \alpha_n^i + \beta_n^i (f_t^n - i_t) + \epsilon_{t+n-1} \\ rx_{t+1}^n &= \alpha_n^x + \beta_n^x (f_t^n - i_t) + \epsilon_{t+1}. \end{aligned}$$

If the expectations hypothesis applied, we would have $\beta_n^i = 1$ and $\beta_n^x = 0$ for all maturities. Estimates for the US bond market are well known (see previous references) and appear to reject the expectations hypothesis. For short maturities, the estimates are $\beta_n^i \approx 0$ and $\beta_n^i \approx 1$. We get the exact opposite from what the expectations hypothesis predicts. Additionally, R^2s are larger and different from zero for the return forecasting equation. As we look at longer maturities, forward rates β_n^i increase and we get some interest forecastability as well.

Does the same conclusion apply for ten years of Brazilian term structure of interest data? Tables 4 and 5 report the results in the sample. I stop at maturity five since the sample is not too long.

[4]There are numerous studies applied to the American market. See Hamburger and Platt (1975), Fama (1976), Shiller, Campbell, and Schoenholtz (1983) Fama (1984), Buser, Karolyi, and Sanders (1996) and Fama (2006).

	a	std(a)	b	std(b)	R-Sq
2	-0.017	*0.008*	1.117	*0.449*	0.209
3	-0.026	*0.010*	0.799	*0.508*	0.108
4	-0.036	*0.011*	1.197	*0.523*	0.210
5	-0.046	*0.015*	2.135	*0.512*	0.389

Notes: Results from the regression $i_{t+n-1} - i_t = \alpha_n + \beta_n(f_t^n - i_t) + \epsilon_{t+n-1}$. Standard deviation calculated from Newey-West estimator with twelve lags (in months). Column "a" refers to estimates of α, column "b" refers to estimates of β.

Table 4: Fama-Bliss Regression: Interest Growth

	a	std(a)	b	std(b)	R-Sq
2	0.017	*0.008*	-0.117	*0.449*	0.003
3	0.024	*0.017*	0.175	*0.677*	0.003
4	0.026	*0.024*	0.372	*0.890*	0.010
5	0.027	*0.030*	0.511	*1.115*	0.012

Notes: Results from the regression $rx_{t+1}^n = \alpha_n + \beta_n(f_t^n - i_t) + \epsilon_{t+1}$. Standard deviation calculated from Newey-West estimator with twelve lags (in months). Column "a" refers to estimates of α, column "b" refers to estimates of β.

Table 5: Fama-Bliss Regression: Excess Returns

Since there is overlap in left-hand variables, I make use of the Newey-West estimator (with twelve lags) of the covariance matrix of residuals, which is consistent for heteroskedasticity and error autocorrelation.

Results look critically different from the US case. In the Brazilian sample, I estimate $\beta_n^i \approx 1$ for $n = 2$, in accordance with the expectations hypothesis. For the other maturities, coefficients are not so close to one - they appear even larger! -, but the difference stays within two standard deviations of it in all cases. We also get a lot of predictability, with R^2s reaching about 20% for two equations and almost 40% for the five year maturity.

As for returns, no signs of predictability. Estimates reported in table 5 start at $\beta_n^x = -0.07$ for $n = 2$ and grow for longer maturities, but they always stay well within one standard deviation from zero. R^2s also do not look good, with the largest one reaching an unimpressive 2%.

These results do not prove the expectations hypothesis holds in the Brazilian market. As we saw, excess returns are large unconditionally, and change with maturity. Plus, regression estimates only say what they say. Our coefficients are close to but not *exactly* zero or one as the hypothesis predicts. More importantly, spreads are not the only variable investors observe. If returns *are* predictable, but only by factors that are *orthogonal* to forward spreads, then the near-zero estimates of β^x's are exactly what we expect them to be. In addition, note from decomposition (3) that the coefficients in the two regressions bear a relationship. A variable that does not forecast cumulative return differences must forecast interest, and vice versa [5]. Therefore, interest predictability in that case would be nothing but the other face of *lack* of return predictability by forward spreads.

[5]For $n = 2$, the connection is clear, since there is only on return term: $f_t^2 = i_{t+1} + rx_{t+1}^2$.

So why do results look so different for the US and the Brazilian cases? In the US case, the explanation for the estimates usually relates to slow adjustment of short-term rates. The β^is grow from zero to one with maturity because it takes a while for interest rates to move in the direction implied by bond prices. Interest will seldom follow a large two-year forward spread (so $\beta_2^i \approx 0$), but will more often adjust after five years (so $\beta_5^i > 0$).

This idea also explains return coefficients. Suppose an investor observes the forward spread, and purchases the n-year bond in any given period. In the following period, suppose further that the interest rate stays the same, and that so does the term structure of forward rates and prices, $p_t^n = p_{t+1}^n$. 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 $\beta^x = 1$ becomes clear. The obvious key step is $p_{t+1}^{n-1} \approx p_t^{n-1}$, which is more likely when the interest rate is unchanged from one period to the next.

And the interest rate is almost *never* unchanged in the case of Brazil. One indication of that is the observation in section 2 that the one-year forward is less correlated with the rest of the curve than the other forward rates. Additionally, table 1 reports an autocorrelation of 0.66 of my interest rate measure in a one-year horizon. Above one, but not that large. The average absolute interest rate growth in the sample is 2.4%, which is economically significant compared to average excess returns (1.5 to 6%), and average forward rates (9 to 12%). We do not even need to appeal to unconditional moments: just compare the variation in the one-year forward plotted in figure 2 to the Fed Funds Rate - stuck at zero during most of the 2010s - and interest variability in Brazil shimmers.

Going back to the example above, if spreads are initially positive $f^n - i_t > 0$, and prices indeed remain the same in the next period, the investor enjoys a positive excess return. The risk, of course, resides in upward shifts of the forward curve. Shift that characterize interest rate hikes. The more these shifts take place, the less confident we are about the investor's strategy success. That appears to be the case in Brazil.

Statistical and sampling uncertainty are two caveats to keep in mind in the interest variability story. After all, I only analyze ten years or so of data, while US dataset often start in the 1950s. If mean interest in Brazil is indeed 9%, then it is true that interest has fluctuated around such average in the sample. Another possibility, however, is that the populational mean interest is different from our sample average and interest rates are indeed *very* persistent. It could be the case, for example, that actual mean interest is 15% per year in Brazil, and we just happened to sample from a period of exceptionally low rates that take a long time to reverse. If that is the case, all we can say is that, at the high frequency at which our dataset conveys relevant information, interest rates have been anything but sluggish around its conditional mean (conditional on low-frequency factors we cannot measure), which is why results from the Fama-Bliss regressions look so different from what literature reports for the US.

I return to the results of the Fama-Bliss regressions in light of my estimated model in section 6.

3.3 Forecasting Returns and the Brazilian Tent

The n-period forward does not forecast the return of the n-period bond. Can we find some variable that does? The natural starting point is to continue to explore the information contained in the term structure. In this subsection, I attempt to find forecastability using as right-hand variables past returns and prices, separately.

Past excess returns do a poor job of forecasting future returns. Table 6 reports the results of a regression of the five-year bond excess return on past excess returns:

$$rx_{t+1}^5 = \alpha + \beta_2 rx_t^2 + \cdots + \beta_{10} rx_t^{10} + \epsilon_{t+1}.$$

Each row of the table corresponds to a different specification in terms of which variables I include on the right-hand side. Numbers reported in italics indicate t statistics. Estimates of the residuals' variance is consistent for heteroskedasticity, and I adjust R^2 for model parameterization. I do not report results using returns from bonds with other maturities as they look very similar.

The first row includes all nine excess return series on the right-hand side of the regression. The large coefficients are an indication of multicollinearity issues, which is expected given that returns are strongly correlated. Still, R^2 reaches only a modest 15.3%. Moreover, no coefficient is statistically significant. Similar results hold when I cut explanatory variables from the equation. All coefficients remain statistically equal to zero, and R^2s , low. No forecastability here.

Next, I regress returns on forward rates, following Cochrane and Piazzesi (2005). Forecasting power is just the same as if we use prices and yields instead of forwards, since they are linear functions of each other. The equation with all forward rates is

$$rx_{t+1}^n = \alpha + \beta_1 f_t^1 + \cdots + \beta_{10} f_t^{10} + \epsilon_{t+1}.$$

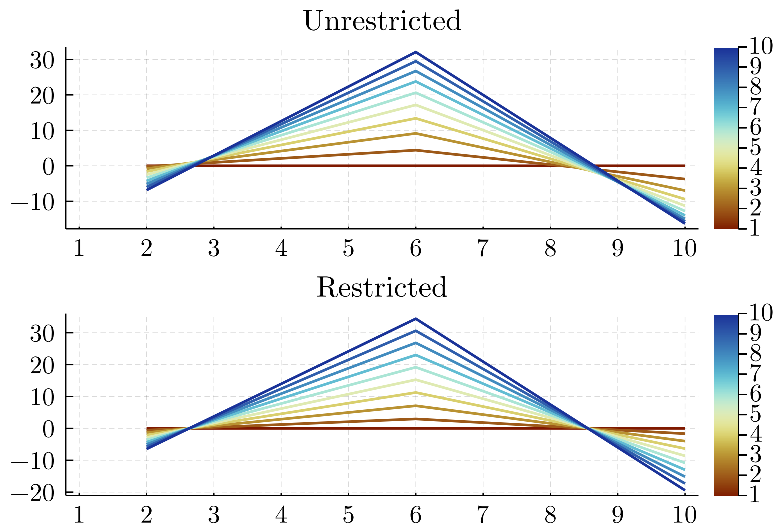
Following Cochrane and Piazzesi, I use 3-month moving averages of forward rates to smooth away short-run noise. More than noise reduction, this procedure appears to capture the effect of lags on excess returns, since forecasting power increases significantly. Table 7 shows results for $n = 5$. I also report results for maturities 2 and 10 years on tables 12 and 13 of the appendix. You can also check results for $n = 2$, 5 and 10 when we use actual forward rates rather than their moving averages in tables 14, 15 and 16 of appendix A.

The regression of returns on all forward rates shows again signs of multicollinearity issues, with absurdly large coefficients. This is also true when I apply the same setting as Cochrane and Piazzesi, using only maturities 1 through 5 on the right-hand side, or when I consider only the five even maturities. The "W"-shape of the coefficients is another indication of collinearity. The fact that we use "data" built from a fitted model can play a part on this, as the Svenson factor model

Const	**rx(1)**	**rx(2)**	**rx(3)**	**rx(4)**	**rx(5)**	**rx(6)**	**rx(7)**	**rx(8)**	**rx(9)**	**rx(10)**	**R-Sq**
0.0		-4.5	-242.7	2018.6	-7882.1	17753.4	-24216.0	19758.3	-8884.4	1695.1	15.3
1.0		*-0.1*	*-0.5*	*0.6*	*-0.7*	*0.7*	*-0.7*	*0.7*	*-0.6*	*0.6*	
0.0		-6.3		14.6		-18.6		10.9		-2.5	13.3
1.1		*-1.2*		*1.2*		*-0.9*		*0.5*		*-0.3*	
0.0		1.2				-0.1				-0.2	9.2
1.1		*0.6*				*-0.1*				*-0.3*	
0.0		1.7				-0.6					9.8
1.2		*1.0*				*-1.5*					
0.0		1.1								-0.3	9.9
1.2		*0.8*								*-1.5*	
0.0						0.7				-0.5	9.3
1.6						*0.7*				*-0.9*	
0.0		-0.6									1.8
1.6		*-0.9*									
0.0						-0.2					6.9
1.8						*-1.6*					
0.0										-0.1	8.3
1.8										*-1.7*	

Notes: t statistic in italic. Results from the OLS regression $rx^5_{t+1} = \alpha + \sum_{n=2}^{10} \beta_n rx^n_t + \epsilon_{t+1}$. Each row presents the estimates for a different set of explanatory variables. R^2 adjusted for model parameterization, and standard deviations calculated from Newey-West estimator with twelve lags (in months).

Table 6: 5y-Bond Return Forecast Regression on Past Returns



Notes: Each curve in the figure plots the loadings of forward rates (maturities 2, 6 and 10 years) on the excess return on bonds of a given maturity: $E_t rx^n_{t+1} = \alpha + \beta' F_t$, $F_t = (f^2_t, f^6_t, f^{10}_t)$. The color bars on the right indicate the n. *Top panel*: coefficients b estimated using unrestricted OLS. *Bottom panel*: coefficients built using the first principal component of $b' F_t$. The figure plots $q\gamma'$, where q is the corresponding eigenvector, and $\gamma = q' b$.

Figure 5: Coefficients of Excess Return Regression on Forward Rates

smooths yield movements that could break multicollinearity. In any case, in using forward rates, our predictive power increases significantly compared to the regression on past returns. R^2 reaches 47% for the specification with all maturities.

All these properties hold when we forecast the return of bonds with other maturities. We get more forecastability and hard-to-interpret coefficients when using a lot of right-hand variables. It turns out we can hold on to the bulk of predictive power - about 40% R^2 in the five-year bond - by using only three forward rates. Maturities 2, 6 and 10 years offer the greatest R^2s for most regressions (I tried many others). Even better, the shape of the coefficients - all of which have high statistical significance -, depicted by the top panel of figure 5, is not so ugly anymore, and closely resembles the "tent" shape of the coefficients found by Cochrane and Piazzesi (2005) for the US market. They suggest that, *ceteris paribus*, expected returns are large when the slope of the first half of the term structure is high and the slope of the second half of the term structured is low. This idea becomes more clear when I provide an algebraic formulation for the return factor. But first, note that we cannot remove more independent variables from the regression without sacrificing additional forecastability (bottom half of table 7). Despite being less parsimonious, I stick to the three rates formulation as it yields the maximum forecastability while involving sensible estimates that match the tent found by Cochrane and Piazzesi.

Let $rx_t = (rx^1_t, rx^2_t, \ldots, rx^{10}_t)'$ be a vector grouping excess returns of bonds with different maturities. Let $F_t = (f^2_t, f^6_t, f^{10}_t)'$ be a vector grouping the rates we use to explain these returns. The unrestricted model I estimate using OLS was $E_t rx_{t+1} = \alpha + \beta F_t$. Note that β is a 10×3 matrix of coefficients. Just like in the US case, the shape of the estimated β with no restrictions cries for a

Const	f(1)	f(2)	f(3)	f(4)	f(5)	f(6)	f(7)	f(8)	f(9)	f(10)	R-Sq
-0.3	6.6	-102.7	1239.9	-9487.7	44198.6	-126913.6	224130.3	-235527.2	134526.3	-32067.9	47.6
-1.8	*1.3*	*-1.1*	*1.1*	*-1.2*	*1.4*	*-1.5*	*1.7*	*-1.9*	*2.0*	*-2.1*	
-0.4	1.0	3.6	-52.2	109.6	-58.4						43.2
-2.6	*0.6*	*0.3*	*-1.3*	*1.7*	*-1.7*						
-0.4	-0.2		-6.7		10.4						28.6
-2.9	*-0.3*		*-2.0*		*2.6*						
-0.3		-1.3		-16.1		106.4		-142.7		56.7	42.7
-2.2		*-0.5*		*-1.0*		*2.0*		*-2.0*		*1.8*	
-0.4		-2.5				17.2				-11.3	39.5
-2.3		*-1.5*				*4.5*				*-2.6*	
-0.3		-1.9								4.5	16.6
-2.3		*-0.9*								*1.7*	
-0.4		-3.2				6.8					29.8
-2.9		*-1.8*				*2.7*					
-0.2						16.1				-13.6	33.1
-1.9						*3.7*				*-3.1*	
-0.1		0.9									6.4
-0.8		*1.4*									
-0.2						2.3					18.1
-2.1						*2.5*					
-0.2										1.9	13.0
-1.8										*2.3*	

Notes: t statistic in italic. Results from the OLS regression $rx^5_{t+1} = \alpha + \sum_{n=1}^{10} \beta_n f^n_t + \epsilon_{t+1}$. I use a three month moving average of foward rates on the right-hand side. Each row presents the estimates for a different set of explanatory variables. R^2 adjusted for model parameterization, and standard deviations calculated from Newey-West estimator with twelve lags (in months).

Table 7: 5y Bond Return Forecast Regression on Forward Rates (3-Month Moving Avg.)

decomposition in which the coefficients for each equation (the rows of β) are a scaled version of a single schedule of parameters. Using symbols, we should be able to find a 10-entry vector q and a 3-entry vector γ such that $\beta \approx q\gamma'$.

To find the pair (q, γ), I follow the procedure adopted by Cochrane and Piazzesi (2009). Selecting (q, γ) is the same as selecting a single factor $\tilde{z}_t = \gamma' f_t$ (the tilde differentiates $\tilde{z}$ from the z I define later) to explain the time-varying part βF_t of expected returns by the equation

$$\beta F_t = q\tilde{z}_t + \eta_t, \tag{4}$$

where η_t is a residual that captures whatever volatility of excess returns not captured by $q\tilde{z}_t$. The first principal component of the series βF_t is a natural candidate to form $\tilde{z}$, as it is built so that the latter inherits the maximum variance from the former (subject to a unitary norm for the weights). Thus, let v_1 be the eigenvector that forms the first principal component, which I now define to be $\tilde{z}$: $\tilde{z}_t = v_1' \beta F_t$. As one can guess from figure 5, $\tilde{z}$ captures over 99% of the total variance of expected returns. By the orthonormality property of eigenvalue decompositions, we can thus write $\beta F_t = v_1 \tilde{z} + \epsilon_t$ for small ϵ_t. Compare that to (4) to see the natural pick $q = v_1$. Finally, to determine γ, return to the definition of $\tilde{z}$:

$$\gamma' F_t = \tilde{z}_t = q' \beta F_t.$$

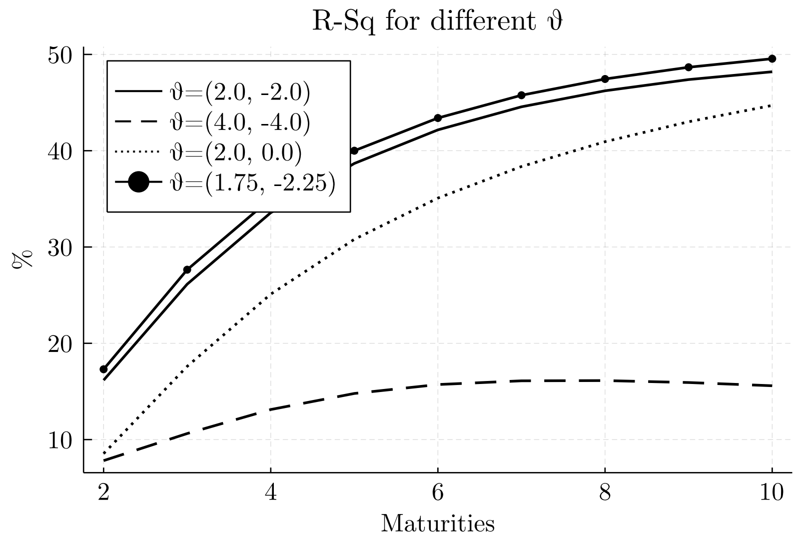
So, $\gamma' = q'\beta$ generates the single factor we are looking for. We have found our pair (q, γ). The bottom panel of figure 5 plots $q\gamma'$. The restricted loadings have shape and magnitude similar to the unrestricted OLS estimates from the top panel.

3.4 Building a Return Factor

The importance of the "tent" shape of the estimated coefficients is that they suggest the existence of a "return factor" that single-handedly explains excess returns almost as well as forward rates. The principal component $\tilde{z}$ calculated above is a natural candidate for such a factor. Consider the following manipulation of the formula that defines it:

$$\begin{aligned}
\tilde{z}_t &= \gamma' F_t \\
&= (-12.2) f_t^2 + (64.4) f_t^6 + (-36.3) f_t^{10} \\
&\approx (-12) f_t^2 + (64) f_t^6 + (-36) f_t^{10} \\
&= 16 f_t^2 + 28(f_t^6 - f_t^2) - 36(f_t^{10} - f_t^6) \\
&\propto f_t^2 + 1.75(f_t^6 - f_t^2) - 2.25(f_t^{10} - f_t^6),
\end{aligned}$$

($\propto$ means "proportional to"). Writing $\tilde{z}_t$ as in the last two lines above is useful to show that it depends on the slopes of the first and second halves of the term structure - as I note in a previous paragraph - as well as on the level of the interest rate. Such dependencies are reminiscent of the unconditional moments studied in section 2. As defined then, the short and long forward spreads $f^{10} - f^5$ and $f^5 - f^1$ capture movements of the two halves of the term structure defined above



Notes: I run regressions of the form $rx_{t+1}^n = \alpha + \beta z(\theta)_t + \epsilon_{t+1}$ for different choices of θ. The table reports the R^2 of each regression. The horizontal axis shows the return maturity n, and each curve represents a different selection of θ.

Figure 6: Predictive Power for the Different Choices of θ

by $f^{10} - f^6$ and $f^6 - f^2$. In table 2, we find that the former is positively correlated with returns, and the latter is negatively correlated. The coefficients above capture such correlations, but now *conditional* on the short-term yield represented by f^2.

Motivated by the expression above, I look for a factor with the following general format:

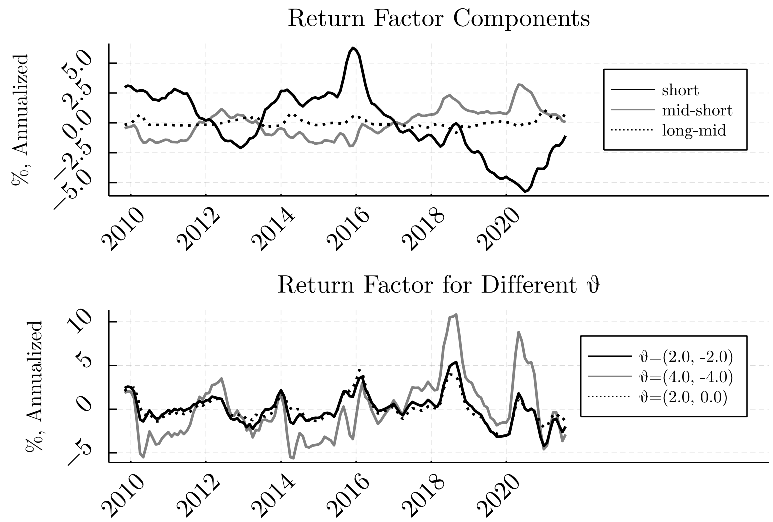
$$\phi z_t = \bar{z} + f_t^2 + \theta_s(f_t^6 - f_t^2) + \theta_l(f_t^{10} - f_t^6), \tag{5}$$

where ϕ is a unimportant positive scalar I later use to set the scale of z_t, and $\bar{z}$ is a scalar I choose so that z_t has zero mean. These two parameters obviously do not affect the factor's ability to explain returns.

The important part is the vector $\theta = (\theta_s, \theta_l)$. Its two elements capture the return's loading on the short slope $(f_t^6 - f_t^2)$ and long slope $(f_t^{10} - f_t^6)$, hence the subscripts. They also govern the importance of the level of interest f_t^2 on the dynamics of z_t relative to the two slope factors. The greater the elements of vector θ in absolute value, the less relevant f_t^2 becomes.

So, how shouold we choose θ? Our re-writting of $\tilde{z}_t$ above suggest $\theta_s = 1.75$ and $\theta_l = -2.25$, and all results throughout the paper remain almost the same if I stick to this choice. But I go one step ahead and set $\theta_s = -\theta_l = 2$ to arrive at a simpler (hence prettier) formula. Figure 6 shows that we sacrifice very little forecastibility by doing such simplication. The figure plots the R^2 for linear regressions $rx_{t+1} = \alpha + \beta z(\theta)_t + \epsilon_{t+1}$ for different choices of θ. The horizontal axis shows return maturity n; each curve represents a different vector θ. The two solid lines correspond to $\theta = (1.75, -2.25)$ (circular markers) and $\theta = (2, -2)$ (no markers). Our simplification does entail losing some predictability, but not too much.

On the other hand, reducing the weight of the long slope by reducing θ_l (dotted line), or that of



Notes: *Top Panel*: time series of each component of the return factor (see (5)): "short" indicates f^2, "mid-short" indicates $f^6 - f^2$, "long-mid" indicates $f^{10} - f^6$. *Bottom Panel*: different instance of $z(\theta)$, as indicated by the legend. I use $\phi = 1$ and set $\bar{z}$ so that the return factor maintains zero sample mean.

Figure 7: Predictive Power for the Different Choices of θ

the short yield f_t^2 by enlarging both θs (dashed line) considerably reduces R^2s . Predictive power is sensible, in particular, to the latter. The next figure 7 helps to investigate why. In its top panel I plot each component of the return factor: f^2 ("short") representing the short-term yield, $f^6 - f^2$ ("mid-short") representing the short slope, and $f^{10} - f^6$ ("long-mid") representing the long slope. In the bottom panel, I plot the same instances of the return factor $z(\theta)_t$ used in the regressions of figure 6. I use $\phi = 1$ in all cases. The solid line in the bottom panel represents the baseline specification of my return factor[6].

The time series of the components show that the two slopes do not span the short yield. In building a return factor, attributing enough weight to f^2 by keeping the θs low is therefore critical to capture various return movements, like the high returns following the 2015-2016 recession and the low returns after 2020[7]. In increasing the absolute value of the θs to 4, the bottom plot makes it clear that we lose the ability to track return in those periods, which is why R^2s drop as they do in figure 6.

I also consider removing the long slope term altogether by setting $\theta_l = 0$. After all, table 1 shows that the long slope is far less volatile than the other terms anyway. Figures 6 and 7 indicate, however, that these few movements matter for predictability. The R^2s of all regressions drop by five to ten percentage points if $\theta_l = 0$. As table 2 portrays, the correlation of the long slope with returns is not the same as that of the short slope. Hence, the long slope adds information. Moreover, having a long slope term in which expected returns are decreasing is consistent with the tent found by

[6]I later set ϕ so that the return factor z has sample variance equal to that of the first principal component of the term structure.

[7]Figure 16 of the appendix plots coefficient estimates for subsamples ending prior to 2021 and verifies stability of my choice of θ both in terms of point estimates and in terms of forecasting power in those subsamples.

Cochrane and Piazzesi (2005) for the US. For these reasons, I stick to $\theta_l = -2$.

Having selected $\theta = (2, -2)$, we need to pick the normalizing constant ϕ. Results are invariant to such choice. I pick ϕ so that the return factor has the same sample variance as the first principal component of the term structure. This choice stabilizes the results from the numerical optimization procedure involved in the estimation below. The first panel of figure 1 plots my return factor, which I henceforth call z.

In all, we can write the time-varying part of the return-forecasting factor for the Brazilian term structure of interest rates as

$$\text{Return Factor} = \text{Short-Term Rate} + 2\,(\text{Short Slope} - \text{Long Slope})\,. \tag{6}$$

In my baseline specification, I use three-month moving averages of the two, six and ten forward rates to build the f^2, $f^6 - f^2$ and $f^{10} - f^6$ series, which serve as my proxies for what the general formula above calls "Short-Term Rate", "Short Slope" and "Long Slope". This need not be case. I also tested the factor's performance when using different proxies. The use of spot rates, different rolling periods for the moving averages and even other choices for short, mid and long rates[8] all yield similar results. Return factors built this way retain most of their predictive power.

3.5 Predictive Power and Principal Components

Table 8 reports the results from return predicting regressions using the baseline return factor z_t and the principal components of the term structure. In it, I set as left-hand variable the five-year return rx_t^5. You can find similar tables (numbered 17 and 18) for maturities 2 and 10 years in the appendix. Each row of the table contains a different specification. Below parameter estimates, I report t statistics in italic font. I use Newey-West estimators of standard deviations (with twelve lags). R^2 is adjusted for the number of model parameters.

Comparing R^2s with the results reported by table 7, we first confirm that we have not lost too much predictive power when rounding up the coefficients. Focusing on the bottom half of the table, we also see that the first three principal components do not explain returns as well as the return factor z. Adjusted R^2 reaches 0.22, compared to 0.38 in the regression with the return factor only. We find comparable predictive power only when including the fourth and fifth principal components.

This is a critical result. It tells us that the information contained in the return factor is *not* spanned by the first three principal components of forward rates. This is only partially due to the fact that I build z using moving averages. By regressing the baseline factor z on the first three principal components, I find an R^2 of about 0.57. When building the return factor with spot rates instead of moving averages, the R^2 increases to 0.62. Far from perfect fit. Forward loadings plotted in figure 4 provide a clue of why this is the case. None of the first three principal components is able to capture the long slope component of the return factor. They all affect mid and long duration

[8] For instance, we can define f^{short} as the cross-sectional average of forward rates with maturities 1 and 2, f^{mid} as the average of maturities 4, 5 and 6, and f^{long} as the average of maturities 9 and 10. Then, using (6), the return factor written $f^{short} + 2((f^{mid} - f^{short}) - (f^{long} - f^{mid}))$ displays similar ability to predict returns as my baseline z_t.

Const	z	PC 1	PC 2	PC 3	PC 4	PC 5	R-Sq
0.0	0.9						38.2
1.8	*3.9*						
0.0	0.9	0.1					38.3
1.7	*3.5*	*0.4*					
0.0	1.0	0.1	-0.4				38.1
1.7	*4.0*	*0.3*	*-0.5*				
0.0	1.0	0.1	-0.5	-0.7			37.7
1.7	*3.8*	*0.2*	*-0.5*	*-0.4*			
0.0		0.5					11.3
1.5		*2.0*					
0.0		0.5	1.9				22.0
1.7		*1.8*	*1.8*				
0.0		0.5	1.8	1.2			21.9
1.7		*1.8*	*1.8*	*0.7*			
0.0		0.6	1.5	1.6	9.5		37.0
1.6		*2.2*	*1.7*	*1.2*	*3.9*		
0.0		0.6	1.5	1.6	9.7	8.1	37.8
1.6		*2.3*	*1.8*	*1.2*	*3.9*	*1.0*	

Notes: Results from the regression $rx^5_{t+1} = \alpha + \beta z_t + \delta' pc_t + \epsilon_{t+1}$, where pc_t groups the term structure principal component (see subsection 2.4). Each row corresponds to a different specification. t statistic in italic. R^2 adjusted for model parameterization, and standard deviations calculated from Newey-West estimator with twelve lags (in months).

Table 8: 5y-Bond Return Forecast Regression on Return Factor and Principal Components

rates in about the same way. Only the fourth component captures such movements of the term structure. Indeed, in regressing the return factor on the fourth principal component as well as the other three, I achieve an R^2 of about 0.985.

Moving now to the first half of table 8, I consider specifications that include as right-hand variables the return factor *and* the level, slope and curve factors. Our ability to predict returns does not increase as we include more principal components. None of their estimated coefficients is statistically different from zero, which indicates that, given the return factors, the traditional level, curve and slope factor do not provide additional relevant information to forecast returns.

These two observations will be critical to provide restrictions to the term structure model, which is where we go next.

4 The Exponential-Affine Model

4.1 Why do we need a model?

Having studied return predictability, our next goal is to decompose forward rates: separate movements due to changes in expected future interest rates from movements in the term premium. It all comes down to predicting the movements of the return factor. If we had the "true" model governing the dynamics of z, we would be done. With it, we could use linear regressions as in subsection 3.5 to forecast the risk premium of bonds with different maturities, and, after that, equation (3) to recover expected interest.

Obviously, we do not yet have a model for the dynamics of z. There are two ways we can proceed. The first is to stipulate a law of motion (such as a VAR) to z and other potential factors and directly estimate the necessary parameters with available data. Then, we can create a discount factor that prices the term structure as precisely as possible. This is the approach proposed by Adrian, Crump, and Moench (2013). It has two main advantages. First, it is computationally simple. With a linear model for the factors and with an exponential-affine model for the stochastic discounting (as the one I present next), we accomplish all we need by running OLS regressions only. The second advantage is that pricing residuals are low: the model does an excellent job in reproducing the observed term structure.

The costs of such convenience, however, can be hard to swallow. The first issue relates to statistical and sampling uncertainty. For example, with highly persistent processes, OLS estimates become volatile and biased. 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-corretion methods have been proposed to address this issue[9]. We cannot rule out the possibility that the data-generating dynamics is also very persistent and that, therefore, OLS estimates are not reliable. The fact that we have few data

[9] See Yamamoto and Kunitomo (1984) for the bias of autoregressive drif estimates. See Phillips and Yu (2005), Tang and Chen (2009) and Bauer, Rudebusch, and Wu (2012) for bias correction in the context of term structure of interest rate models.

points compared to studies of other markets only makes matters worse. A second and even more problematic downside of direct estimation is that it heavily relies on enough freedom in the building of the discount factor. In the case of the exponential-affine model, for example, the market price of risk process contains a total $M + M^2$ free parameters to be estimated, where M is the number of factors. Imposing restrictions - such as having a single factor govern expected returns - significantly compromises model fitting.

In this paper, I follow a different approach, pursued by Cochrane and Piazzesi (2009). I infer the law of motion of the factors by first estimating the risk-neutral dynamics *from the cross section* of yields. Under the risk-neutral dynamics, forward rates vary only due to variation in expected interest. The fact that we observe in any given point in time forward rates of different maturities means that we observe expected interest rate in different time horizons. The estimation step explores this property of the risk-neutral dynamics to pin down the law of motion of the model's factors along with their impact on the interest rate. The affine model then provides a link between risk-neutral and the data-generating dynamics through market prices of risk, which I calibrate to reproduce the findings of the previous section. The drawback of this indirect approach is that the estimation is non-linear and thus numerically challenging.

I now present the equations of the model and then provide more details of the estimation procedure.

4.2 The Exponential-Affine Model

Following Ang and Piazzesi (2003) and Ang, Dong, and Piazzesi (2007), I consider the homoskedastic, discrete time version of the multifactor exponential-affine model presented by Duffie and Kan (1996). In stating the model, I follow the same notation as in section 2. Definitions of forward rates and excess returns extend to the model.

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

$$X_t = \mu + \Phi X_{t-1} + e_t, \tag{7}$$

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 a $n-1$ 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

$$i_t = -p_t^1. \tag{8}$$

So, the pricing equation above implies

$$0 = \log E_t \exp\left\{m_{t+1} + i_t + rx_{t+1}^n\right\}, \qquad rx_{t+1}^1 = 0 \text{ a.s.} \tag{9}$$

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

$$m_{t+1} = -\delta_0 - \delta_1' X_t - (1/2)\lambda_t' \Sigma \lambda_t - \lambda_t' e_{t+1} \tag{10}$$

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:

$$\lambda_t = \lambda_0 + \lambda_1 X_t, \tag{11}$$

where λ_1 is square matrix.

Equation (8) gives the solution for the interest rate:

$$i_t = \delta_0 + \delta_1' X_t. \tag{12}$$

Replacing the discount (10) on the general solution (9) yields

$$E_t rx_{t+1}^n = -(1/2)\text{var}_t(rx_{t+1}^n) + \text{cov}_t(rx_{t+1}^n, e_{t+1}')\lambda_t. \tag{13}$$

The covariance term represents the amount of risk of each bond, or its "beta"[10], which is why λ_t is referred to as the market price of risk. Given the solution below, the conditional variance and covariance terms 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

[10] 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 homoskedasticity 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" beta, and $\Sigma\lambda_t$ becomes the market price of risk.

same is true for forward rates:

$$p_t^n = A_n + B_n' X_t \tag{14}$$
$$f_t^n = A_n^f + B_n^{f\prime} X_t. \tag{15}$$

The relationship between forward rates and prices (1) implies $A_n^f = A_{n-1} - A_n$ and $B_n^f = B_{n-1} - B_n$. Using our definition (8) of the nominal risk-free rate and the discount factor (10), we also find that $-A_1 = A_1^f = \delta_0$ and $-B_1 = B_1^f = \delta_1$. By replacing (14) on the definition of excess returns (2), and that on the pricing equation (13), we find expressions that allow for the computation of the solution coefficients recursively:

$$\begin{aligned} A_n^f + B_{n-1}' \mu^* &= -(1/2) B_{n-1}' \Sigma B_{n-1} + \delta_0 \\ B_n' &= -\delta_1' + B_{n-1}' \Phi^* = -\delta_1' (I + \Phi^* + \Phi^{*2} + \cdots + \Phi^{*n-1}) \\ B_n^{f\prime} &= \delta_1' \Phi^{*n-1}. \end{aligned} \tag{16}$$

In the equations above, I define

$$\begin{aligned} \mu^* &= \mu - \Sigma \lambda_0 \\ \Phi^* &= \Phi - \Sigma \lambda_1. \end{aligned} \tag{17}$$

These coefficients characterize the *risk-neutral* dynamics of the factors X_t. To see this, replace the computed coefficients in (16) in (15):

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

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, (9) 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 (7). 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).

5 Restrictions and Estimation

5.1 Factor Choice and Estimation Strategy

I consider a four factor model. The first factor is the return-forecasting factor z. The next three factors are the first three principal components of the term structure, to which I refer as level, slope and curve factors from now on. I normalize all factors to have zero mean.

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

These are the factors figure 1 plots. Postulating economic variables that drive yields from theory is beyond the scope of this paper. As I argue in section 2, when that is the case the eigenvalue decomposition of yields offers natural factor candidates since we can interpret innovations in terms of intuitive and orthogonal movements in the term structure.

On a high level, the estimation of the model is simple. I select μ^*, Φ^* and $\delta = (\delta_0, \delta_1)$ to minimize pricing errors using equation (15), and then λ to reproduce a linear regression model of expected returns using (13). With Φ^* and λ, I use (17) to infer the data-generating dynamics.

Our conclusions of section 3 are suggestive of the format of the linear model to be used in the second step of the procedure, given our choice of factors. We conclude that, given the return factor, the other three factors should not predict returns (table 8). It suggests, therefore, a model of the format $E_t rx^n_{t+1} = \alpha + \beta z_t$. Let $rx_t = (rx^2_t, \ldots, rx^{10}_t)'$. In the second step of the procedure we therefore attempt to reproduce

$$E_t rx_{t+1} = a + bz_t = -(1/2)B'\Sigma B + B'\Sigma(\lambda_0 + \lambda_1 X_t), \tag{18}$$

where $B = [B_1 \ldots B_9]$. The second equality follows from (13) evaluated at the model's solution. Condition (18) implies that the columns of λ_1 corresponding to the level, slope and curve factor must be zero. I thus constrain twelve of the sixteen parameters in λ_1 to be zero and significantly reduce the model's parameterization.

The restriction I impose above says that market prices of risks vary over time only due to variation in z. This is an immediate consequence of the fact that we can explain expected returns with a single variable. Cochrane and Piazzesi (2009) use a similar constraint in their estimation. Since we work with a conditionally homoskedastc model, restricting the price of risk is enough to ensure that only z forecasts risk premia in the model, which is the feature of the data we would like to reproduce.

5.2 Cross-sectional Covariances and Risk Factor Selection

I build a model in which market prices of risk vary only with the return factor. This is not to say that these prices of risk compensate investors for exposure to return factor shocks only. Asking which factors drive expected returns is not the same as asking exposure to which shocks this expected return compensates investors for. Analytically, having λ_t depend only on X_t does not mean that all of its entries but the first are equal to zero.

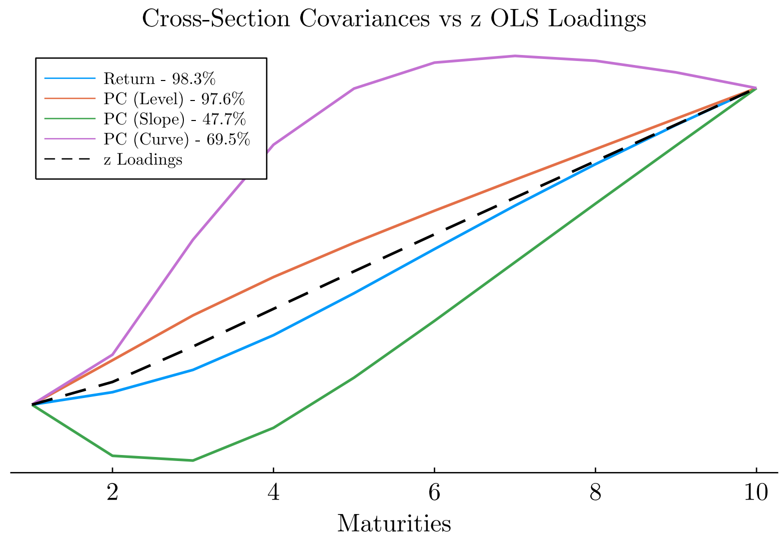
In the absence of a structural macroeconomic model, we have no theoretical basis to come up with constraints on how each of the four shocks in our model is priced. Nevertheless, the shape of the *cross-sectional covariance structure* of each shock (or how the return of bonds with different maturities covaries with them) can provide clues of which of them are more likely to be priced. To see this, re-write (18):

$$
\begin{aligned}
E_t rx_{t+1} &= a + [b]z_t \\
&= (.) + \left\{[\mathrm{cov}(rx_{t+1}, e_t^z)]\lambda_1^z + [\mathrm{cov}(rx_{t+1}, e_t^l)]\lambda_1^l + [\mathrm{cov}(rx_{t+1}, e_t^s)]\lambda_1^s + [\mathrm{cov}(rx_{t+1}, e_t^c)]\lambda_1^c\right\} z_t
\end{aligned}
$$

I use brackets to highlight the fact that the terms inside are vectors. What the expressions tells us is that the vector of loadings b is a linear combination of the term structure of covariances with each shock, each represented by a $\mathrm{cov}(rx_{t+1}, .)$ term. The elements of the first column of λ_1 (written as λ_1^z for the return factor z, and so on) provide the weights. If the shape of a given covariance schedule is too different from vector b, then the corresponding $\lambda_1^{(.)}$ must be zero or small compared to the others. In the US case, for instance, Cochrane and Piazzesi (2009) show that only the schedule of covariances with level shocks line well with their z loadings. The remaining covariance vectors have completely different shapes. They rely on this observation to add the restriction that only level shocks have a non-zero market price of risk.

Sadly, in the Brazilian case the same analysis leads to two viable candidates for risk factor. Figure 8 plots all relevant vectors. The dashed line corresponds to b, calculated using linear regressions as in section 3. To calculate the covariance schedules, I run OLS estimates of (7), compute residuals and calculate their covariance with each of the excess return series. I multiply each covariance schedule by a constant selected to ensure that the last point on the right (maturity 10) is the same. I also report on the figure's legend the R^2 of OLS regressions of b on each covariance vector individually, without adding constants.

Expected return loadings increase linearly with bond maturity. As R^2s confirm, only the return shock and level shock covariance schedules display such linear shape. We cannot, however, rule out either of them based on the plot only. Given that both factors are good enough on their own to generate the right pattern of return loadings (R^2 superior to 0.97), I consider only a specification in which there is a single risk factor. In my preferred specification, I assume the level factor to be the only risk factor, for three reasons. First, I find better model fit and more stable results. Assuming return shock is priced instead leads to highly explosive dynamics under the data-generating probability measure. And we know the world does not diverge. Second, when building the return



Notes: The dashed line is the loading of excess returns of different bonds on the return factor z. The horizontal axis represents bond maturity. Each solid line plots the covariance of the indicated return with a different model shock. All curves are re-scaled to end at the same point on the right. To build time series for the shocks e, I run OLS estimates of the parameters in (7) and compute residuals. The legend contains R^2 of individual linear regressions of loadings (dashed line) on each covariance schedule.

Figure 8: Cross-sectional Covariance Schedule *vs* Expected Returns

factor using forward spread rather than levels, the covariance schedule completely changes its shape, and that does not happen to the schedule of covariances with level shocks. Third, as I wrote before, having only level shocks being priced is consistent with results for the US bond market.

Assuming only levels shocks are priced means that $\lambda_1^z = \lambda_1^s = \lambda_1^c = 0$ in the equation above. I thus reduce λ_1 to a single parameter to be estimated. In addition, only the second entry of λ_0 - call it λ_0^l - is different from zero. We have reduced a twenty-parameter model for the market price of risk to a two-parameter model.

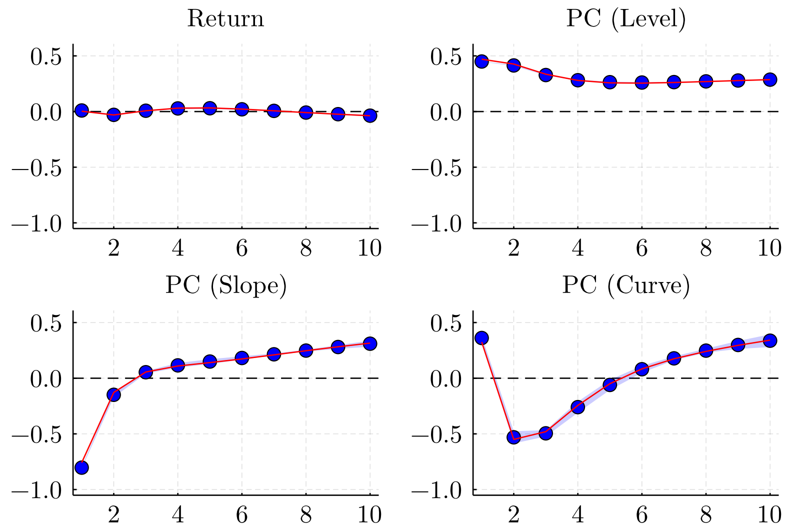
5.3 Estimation

I use the OLS-estimated series of shocks to estimate their covariance matrix Σ. Let Σ_l be the second column of Σ.

In building the factors, I have demeaned each time series, and I now assume $\mu = 0$. That is, all factors have zero unconditional mean. In essence, I exchange uncertainty about μ for uncertainty about $E(X_t)$. Nothing guarantees that factors have indeed fluctuated around their unconditional means over the last ten years. Setting $\mu = 0$ yields:

$$\mu^* = -\Sigma_l \lambda_0^l \tag{19}$$

I group parameters δ_0, δ_1, λ_0^l, Φ^* in a single vector θ. Notice that, given (19), these are all the parameters that we need in order to calculate the forward rates implied by the affine model. Coefficients A_n^f and B_n^f do not depend on data-generating dynamics Φ or the price of risk parameter



Notes: Each figure above plots OLS estimated coefficients of linear regressions $f_t^n = a + b'X_t + \epsilon_t$ of forward rates on model factors (blue markers) and the coefficients $B^f(\theta)$ implied by the exponential affine model and evaluated at the optimal solution (red curves).

Figure 9: OLS *vs* Model-Implied Linear Coefficients

λ_1. So, define $\hat{f}^n(\theta; X)$ as the value taken by the forward rate of maturity n according to solution (15), when we use the set of parameters θ and X is the current vector of factors. I minimize the mean squared error implied by the model to estimate θ:

$$\text{Min}_\theta \qquad \left[\frac{1}{TN} \sum_t \sum_{n=1}^{N} \left(f_t^n - \hat{f}^n(\theta; X_t) \right)^2 \right]^{\frac{1}{2}} \qquad \text{s.t.} \quad (19),$$

where $N = 10$, using Newton-type search optimization algorithms. The procedure achieves a mean squared error of 32 bps (basis points) and a mean absolute error of 25 bps. Unrestricted OLS estimates of the regression $f_t^n = a_n + b_n' X_t + \epsilon$ give the coefficients that minimize squared errors equation by equation in the class of linear models. Therefore, we can use it as a benchmark for the restricted environment. OLS yields a mean squared error of 11 bps and a mean absolute error of 8 bps. Despite the difference, figure 9 shows that the exponential affine model has enough flexibility to almost perfectly replicate OLS. In it, the blue markers represent the OLS coefficients, and the red curves represents the sequence of B_f coefficients implied by the model, when evaluated at the optimal θ.

The first step of the optimization is complete. The second step involves estimation of the market price of risk parameter λ_1. We reduce its sixteen entries to a single non-zero one, λ_1^l, the second element of the first column. In light of equation (18), our goal is to have the time-varying component of risk premia depend on z_t in the model ($B'\Sigma\lambda_1 X_t = B'\Sigma_l \lambda_1^l z_t$) like we estimate it to depend in the data (the b). So, I estimate $E_t rx_{t+1} = a + bz_t$ by OLS, equation by equation, like in section 3. I then regress (without an intercept) the estimated $\hat{b}$ on the model-implied betas $B'\Sigma_l$. Note that these are both one-dimensional vectors. Referring again to figure 8, I set λ_1^l by regressing the dark

	Return	PC (Level)	PC (Slope)	PC (Curve)	St Dev %
Return	1.00	0.39	0.68	0.17	6.08
PC (Level)		1.00	0.19	-0.12	4.94
PC (Slope)			1.00	0.18	1.35
PC (Curve)				1.00	0.51

Notes: I estimate (7) by OLS and compute residuals. The table above reports the Pearson correlation matrix of such residuals and, on the right column, their sample standard deviation.

Table 9: Estimated Shock Correlation Matrix and Volatility

dashed curve on the red curve[11].

Finally, having Φ^*, Σ and λ_1^l I recover the true model dynamics from (17). The estimation is complete.

5.4 Dynamics and the Market Price of Risk

I now describe estimated dynamics under the risk-neutral and data-generating probability measures.

Table 9 reports the correlation between factor shocks, as well as their standard deviation, all calculated from my estimate of Σ. Figure 17 of the appendix plots the underlying series. The standard deviation of return shocks is one percentage point higher than that of levels shocks. Both are significantly more volatile than slope and curve shocks. Having said that, and given the assumption that only level shocks are priced, the 0.39 correlation coefficient between return and level factor shocks is the most relevant one from the table.

The first step of the estimation procedure yields $\lambda_0^l = -4.15$. The negative sign is critical for the interpretation of factor shocks. Since model and sample factor means are zero, a negative λ_0^l indicates that, on average, the second element of λ_t is negative. From (13), that means that - on average - any asset or portfolio with a payoff positively correlated with the level shock receives a penalty in terms of expected return. We infer from this point that level shocks capture jumps of marginal utility in the same direction. An asset that pays more when a positive level shock hits is less risky than assets that do the opposite, and provides investors with *insurance* against spikes in marginal utility.

Nominal bonds as we consider in this paper do the exact opposite. Model parameters B (unreported, see equation (14)) are negative in their second entry: positive level shocks coincide with negative price jumps. Bonds, therefore, are risky investments in that their payoff is lower when marginal utility - the level factor - is higher. These shocks that drive down prices and lead to low realized returns also tend to come accompanied by increases in expected return in the following period. More intuitively, periods of low consumption are characterized most of the time by higher risk premia.

[11] I compute the covariance schedules of figure 8 using the sample covariance of rx and estimated shocks e. In the second step of the estimation here described, I use the model-implied covariance $B'\Sigma$. Since the estimated B closely matches the data, empirical and model-implied covariance schedules are similar.

	Const.	Return	PC (Level)	PC (Slope)	PC (Curve)	|Eig|
Return	0.00	0.67	0.38	-0.22	0.74	1.02
PC (Level)	0.01	-0.02	0.93	0.45	-0.14	0.87
PC (Slope)	0.00	0.01	0.05	0.51	0.87	0.41
PC (Curve)	-0.00	-0.06	0.08	0.14	0.52	0.41

Notes: I estimate risk-neutral dynamics to minimize pricing errors, and market-price of risk parameters to reproduce the estimated linear regression of risk premia on the return forecasting factor z. I then use (17) to compute the model's dynamics under the data-generating probability measure. The table describes factor dynamics under the *risk-neutral* probability measure. The first column reports μ^*. The next four columns report Φ^*, and the last column reports the absolute value of its eigenvalues.

Table 10: Estimated Factor Dynamics under the Risk-Neutral Probability Measure (μ^*, Φ^* and Eigenvalues)

	Const.	Return	PC (Level)	PC (Slope)	PC (Curve)	|Eig|
Return	0.00	0.34	0.38	-0.22	0.74	0.89
PC (Level)	0.00	-0.72	0.93	0.45	-0.14	0.76
PC (Slope)	0.00	-0.03	0.05	0.51	0.87	0.76
PC (Curve)	0.00	-0.06	0.08	0.14	0.52	0.19

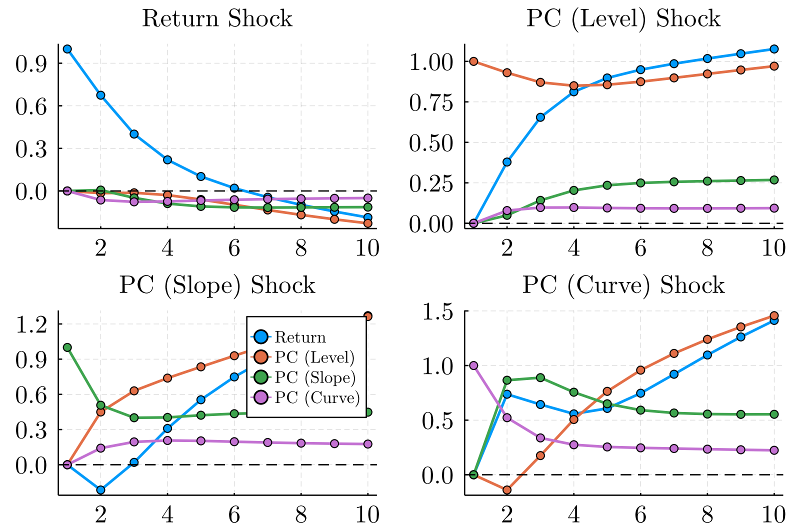
Notes: I estimate risk-neutral dynamics to minimize pricing errors, and market-price of risk parameters to reproduce the estimated linear regression of risk premia on the return forecasting factor z. I then use (17) to compute the model's dynamics under the data-generating probability measure. The table describes factor dynamics under the *data-generating* probability measure. The first column reports μ. The next four columns report Φ, and the last column reports the absolute value of its eigenvalues.

Table 11: Estimated Factor Dynamics under the Data-Generating Probability Measure (μ, Φ and Eigenvalues)

The next two tables, 10 and 11, describe factor dynamics. Table 10 focuses on risk-neutral dynamics, and reports μ^* (first column) and Φ^* (next four columns). The final column reports the absolute value of each eigenvalue, in increasing order. Table 11 has the same format, but reports the actual dynamics.

The constraint that only the return factor affects the market price of risk implies that only the first column of Φ changes when we build the risk-neutral Φ^*. In words, under the risk-neutral measure, the loadings of each factor on the previous value of the return factor changes. The remaining loadings do not. Moreover, from the reported values, one can argue that only two coefficient changes are economically significant: the first and second elements of the first column, which go from $(0.34, -0.72)$ to $(0.67, -0.02)$. Imposing constraints on λ_1 has certainly led to larger pricing errors. A clear connection between data-generating and risk-neutral dynamics is where it pays off.

We can see from the tables that changing the probability measure makes factors *more* persistent. This result differs from the findings of Cochrane and Piazzesi (2009) for the US. While stationary dynamics under the data-generating measure is desirable - again: the world does not diverge -, a near-unitary root for our factors under the risk-neutral probability measure is to be expected. Referring back to (3), we know that the first principal component of the term structure of interest



Notes: Each panel plots the impulse response functions of the model's factors to the shocks indicated by their title, under the risk-neutral probability measure.

Figure 10: Impulse Response Function under the Risk-Neutral Probability Measure

soaks up about 93% of the total variance of forward rates. With risk-neutral pricing, forward rate variation simply reflects variation in expected future interest. Therefore, "level" shocks affect expected interest more or less uniformly in the time dimension, which is only possible if they are persistent. This is why we must have an axis rotation corresponding to a near unity eigenvalue in of matrix Φ^*.

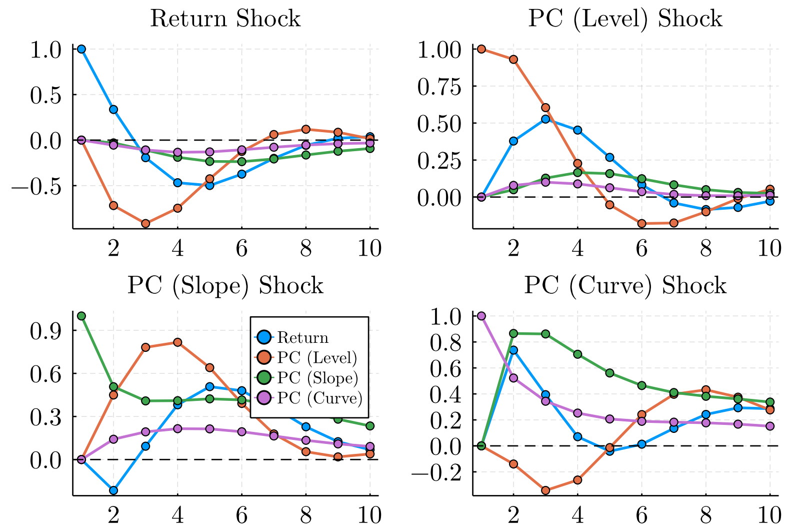
What is not necessarily to be expected, and hence constitutes a more interesting result, is that shocks to model factors are also persistent under the data-generating measure. The largest eigenvalue of Φ is 0.89. As means of comparison, OLS regression of the VAR (7) yields an estimate for Φ with the largest eigenvalue being 0.57 in absolute value (see table 19 in the appendix). Our procedure to estimate the data-generating dynamics indirectly yields wildly different results than one would find by running OLS. In section 6, I consider how the additional persistence of shocks affect forward rate decompositions.

Figures 10 and 11 translate hard-to-read autoregressive matrices Φ^* and Φ to their equivalent moving average represention s. Each panel plots factors' impulse response functions to innovations on one of the four shocks. Figure 10 displays risk-neutral dynamics, figure 11 displays data-generating dynamics. It is easier to start by reading the data-generating dynamics. The top panels of figure 11 reveal two important movement patterns I highlight.

1. Rising level factor pulls up the return factor;

2. Rising return factor pushes down the level factor.

Linearity implies the reverse statements hold as well. The dynamics we estimate suggests that, whenever the level factor - or marginal utility - increases, that tends to be accompanied by rising risk premium in the following years (coefficient 0.38 of table 11). This is what we expect: during



Notes: Each panel plots the impulse response functions of the model's factors to the shocks indicated by their title, under the data-generating probability measure.

Figure 11: Impulse Response Function under the Data-Generating Probability Measure

recessions and periods of low consumption, investors demand larger returns in order to hold (risky, as we saw) bonds. Factor level pulls up factor z.

If the first statement above portrays how the economy enters periods of low consumption, the second statement give us clues of how these periods come to an end. The top-left panel of figure 11 shows that higher than average return factor is followed by movements of the level factor in the opposite direction. High risk premium leads to declining marginal utility. This movement is captured especially by the first column, second row element of matrix Φ (-0.72). Return factor z pushes the level factor in the opposite direction.

And it is precisely the change of *this* coefficient that leads to the larger persistence of the risk-neutral dynamics we estimate. In the case of matrix Φ^*, the same coefficient is -0.02, which means that the "ability" of the return factor of "pushing down" and stabilizing marginal utility is, to a large extent, gone. Indeed, the top left panel of figure 10 shows that, under the risk-neutral probability measure, movements in the return factor come roughly unaccompanied by movements in the other three factors. Compare that to the same panel of figure 11 to see the clear difference. In practice, this means that when the level factor moves - top right panel - the resulting variation in risk premium is not enough to stabilize it. Hence, the protracted response of marginal utility. Under the risk-neutral measure, booms and recessions last longer.

Finally, I also point out that, under the data-generating probability measure, all four shocks in the model lead to movements in the return factor and, hence, in risk premia. The timing and shape of such movements is heterogenous, however, which leads to also heterogeneous profiles of term premia in the maturity dimension, as we see in section 6.

5.5 *Discussion*

We cannot price assets by discounting expected payoffs at the risk-free rate because payoffs have different value to investors in different states of nature. Risk-neutral probability measures simply re-weight probabilities of different states of nature so that the states investors "care" the most about - at least in terms of their utility - become more likely to happen. We can then forget about marginal utility altogether and price payoff streams by simply discounting them at the risk-free rate.

The statements we make about the risk-neutral measure when comparing figures 10 and 11 are, therefore, statements about investors' marginal utility. Our results imply that the marginal investor is particularly sensible to sequences of shocks that lead to long periods of high yields across the term structure. Looking back on the sample, the 2015-16 interval calls attention as a period of remarkably high yields of all maturities. During these years, the Brazilian economy experienced a large recession, with soaring inflation and unemployment. Our results thus support an intuitive narrative, in which investors "fear" - at least to the extent that "fear" can be read "discount less" - realizations of uncertainty that lead to prolonged periods similar to the 2015-16 recession.

Pointing out which variables best explain the link between recession and marginal utility is hard. For instance, unemployment skyrocketed during the COVID-19 pandemic as well, but during the same period risk premium fell considerably. Inflation was high during the recession and low during the pandemic. So maybe inflation is the main driver of marginal utility? Perhaps, but the answer to that question will ultimately require additional data and, hopefully, an underlying theory that goes beyond the scope of this article.

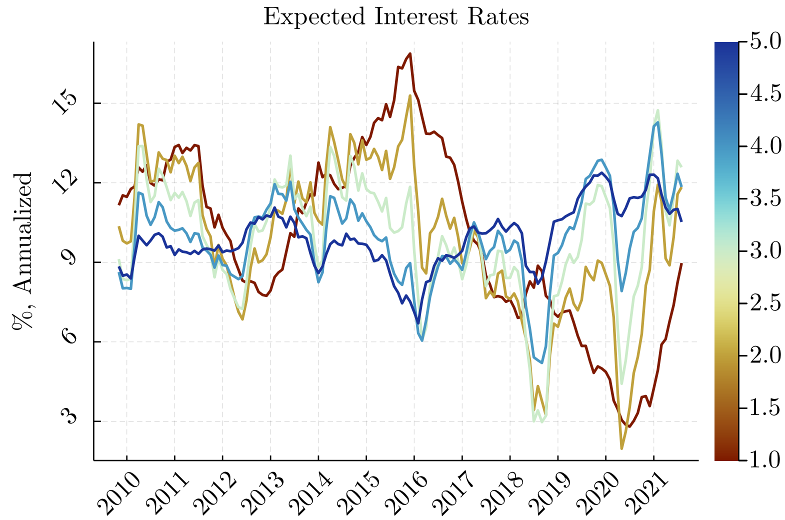
6 Results

6.1 *Yield Decomposition and the 2015-16 and 2019-20 Episodes*

Having estimated its parameters, we can use the model to project the path of the factors (markers in figure 1) and the path of interest rates. Expected interest is given by $E_t i_{t+n} = \delta_0 + \delta_1' \Phi^n X_t$. We can also compute the term premia as the difference between model-implied forward rates and expected interest.

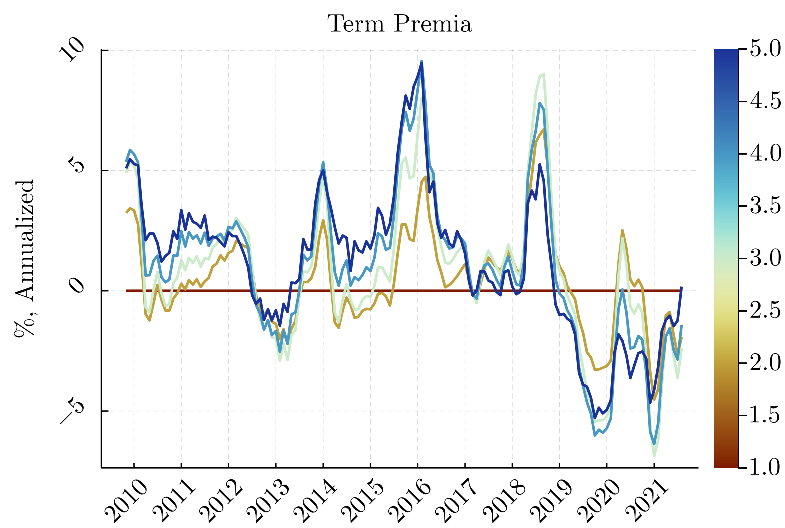
I present the decomposition in figures 12 and 13. The former contains expected interest $E_t i_{t+n-1}$, the latter contains the term premium $f_t^n - E_t i_{t+n-1}$. In each figure, lines of different colors indicate different maturities. Lines of the same color across figures correspond to the same decomposition. For instance, the blue curves in each plot combine to form the five-year forward yield. I stop at the five-year maturity to avoid making the graph too crowded.

Our model produces times series of expected interest rates with volatility comparable to that of forward rates. The expected interest rate four years in the future ranges from 6 to 13%; the five-year forward from 7 to 17%. Such volatility follows from the large eigenvalue present in the factors' dynamics Φ - in particular the persistence of the level factor. Innovations take a long time to dissipate, and greatly affect our expectations of short-term rates over horizons of five to ten years.



Notes: The figure plots the model-implied expected interest rate. Section 5 describes the steps of the estimation of the exponential-affine model. I use estimated parameters to compute expected interest using $E_t i_{t+n} = \delta_0 + \delta_1' \Phi^n X_t$. Each curve corresponds to a different maturity n. The color bar on the right indicates n.

Figure 12: Model-Implied Expected Interest Rates $E_t i_{t+n-1}$



Notes: The figure plots the model-implied term premia. Section 5 describes the steps of the estimation of the exponential-affine model. I use estimated parameters to compute expected interest using $E_t i_{t+n} = \delta_0 + \delta_1' \Phi^n X_t$. The term premia is difference between model-implied forward rates and expected interest rate. Each curve corresponds to a different maturity n. The color bar on the right indicates n.

Figure 13: Model-Implied Term Premia $f_t^n - E_t i_{t+n-1}$

Interest persistence, however, does not mean positive autocorrelation at all lags. On the contrary, as figure 11 suggests, factors tend to fluctuate around their unconditional average before stabilizing[12]. Interest rates inherit the same pattern. Times of high interest are not necessarily times of high expected interest in the long term. Our estimates indicate that expected interest four years in the future are higher in the final third of the sample, when short-term rates were lower, than in the previous two. We can therefore conclude that forward rates are poor indicators of expected interest.

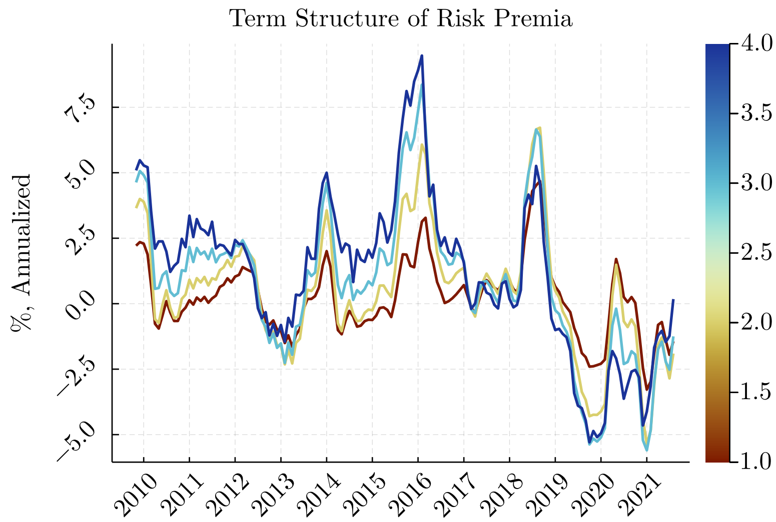
What about term premia? Figure 13 shows that the term premium is large on average and volatile, with both moments increasing in bond maturity. As the other face of the fluctuation of expected interest around its average, the term premium of long-term bonds is higher when interest is higher and vice-versa. Our calculations suggest that the large upswing of the term structure during the 2015-2016 recession and its downswing in 2019-2020 were *both* caused not by a change in the perception of future interest, but rather by a change in bond premia, or investors' willingness to hold public debt.

We can also describe 2015-2016 and 2019-2020 episodes in terms of the two movement patterns I highlight in section 5. In 2015-2016, the level and the return factors peak following positive innovations to each of them (figures 1 and 17, latter in the appendix). As we saw, a rising return factor predicts a lower level factor. Indeed, level factor declines starting in 2016. This is consistent with a lower predicted short-term rate in the coming years, which is exactly what figure 12 claims. While the short-term interest reaches its *maximum* by the end of 2015, projected short-term interest in the four-year horizon reaches its *minimum*. High forward rates were therefore capturing mostly a large risk premium, which is exactly what figure 13 claims.

We can tell a similar story for the fall of term structure rates in 2019 and 2020. The sequence of negative shocks to the level factor that starts in 2018 leads to low rates and - as *per* the first movement pattern I highlight - a low return factor. A below-average return factor in turn - as *per* the second movement pattern - predicts a rising level factor and, by extension, a rising short-term rate. The low forward rate and high expected interest imply a low term premium, which figure 13 verifies.

By the time I write this article, the return factor remains below its unconditional mean. The other three factors are close to their mean. The level factor, in particular, has increased back to zero, following a decline in the return factor (as our dynamics would have predicted), but also due to a large positive innovation. Figure 12 shows that such reversal has had little impact on projected long-term interest in three year and over. The fact that they remain above average we can attribute to the path followed by the return factor, which remains *below* average. In fact, note that the forecast for the future path of all four factors basically replicates the impulse-response function to a return shock, as depicted by figure 11.

[12] Matrix Φ contains complex eigenvalues, which capture the "wave" movements of the dynamic system.



Notes: The plot represents the decompositon of the term premium of the five-year duration bond over time. Each curve corresponds to the cumulative sum of the different components of the term premium, as indicated by expression (3). The color bar indicates the term farther in the future in terms of *item*. For example, the red curve represents $E_t\left(rx^5_{t+1} - rx^4_{t+1}\right)$, the next one (yellow) represents the same plus $E_t\left(rx^4_{t+2} - rx^3_{t+2}\right)$, and so on. I calculate expectations using the estimated model parameters.

Figure 14: Decomposition of the 5y Term Premium - Cumulative Sum of Risk Premia

6.2 Term Structure of Risk Premia and Revisiting Fama-Bliss

We can also decompose term premia. As expression (3) shows, the term premium is the sum of the differences of risk-premium of consecutive maturity bonds over time:

$$\text{Term Premium}^n_t = E_t\left(rx^n_{t+1} - rx^{n-1}_{t+1}\right) + E_t\left(rx^{n-1}_{t+2} - rx^{n-2}_{t+2}\right) + \cdots + E_t(rx^2_{t+n-1}).$$

Having a model to predict the return factor z, we can predict excess returns rx, and thus compute the different components above. The question we try to answer is where is risk located on time, or what is the *term structure of risk premia*?

Cochrane and Piazzesi (2009) estimate that, in the US market case, the term structure of risk premia is non-trivial in the sense that a large risk premium today is no indication of large premia in the future and vice-versa. The authors argue, for this reason, that there is no sense in referring to a "forward premium". Each forward rate will reflect a different term premium, and term premia of different maturities need not be proportional to each other or even share the same sign.

Figure 13 suggests that the same result holds for the Brazilian market as well. In fact, at the time I write, the risk premium is low, as indicated by the negative return factor in figure 1. Hence, the negative term premium *of short term bonds*. The model forecasts, however, that returns will be high in four years from the current date (*expected* returns will be high in three years, red markers in figure 13). That should reflect *only* on the term premia of bonds with duration five years and over. This is exactly what figure 13 shows. The term premium of the five-year bond is considerably larger than that of bonds with lower maturity. Things add up.

Figure 14 explains the term structure of risk premia in a different way. Consider only the forward premium of the five-year duration bond. Equation 3 says we can decompose it in four risk-premium components. The figure plots the *cumulative* sum of each of these components. For this plot, the color bar indicates time, not maturity. The red curve, corresponding to $t+1$ as indicated by the color bar, represents $E_t\left(rx^5_{t+1} - rx^4_{t+1}\right)$. The next curve ($t+2$, yellow line) represents $E_t\left(rx^5_{t+1} - rx^4_{t+1}\right)$ (the red curve) *plus* $E_t\left(rx^4_{t+2} - rx^3_{t+2}\right)$. And so on. The blue curve sums up all terms of the decomposition, and thus represents the term premium itself.

Consider the last point of the data set. I argue above that the term premium of the five-year bond is due mostly to risk premium "located" three years from now, expected to realize in the fourth year. Figure 14 confirms that assertion, as the blue curve - the only one capturing risk-premium for the four-year return - jumps above the other components.

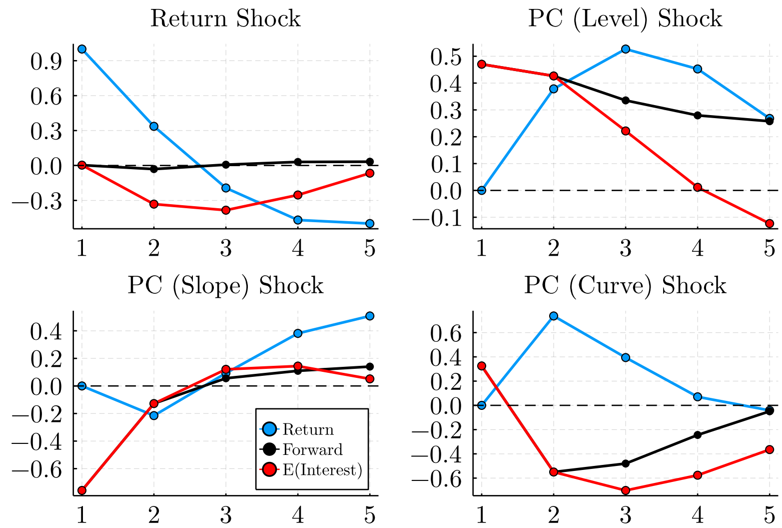
The figure also shows that the principle of a "non-trivial" term structure of risk premia holds in other moments of time. If risk premium was given by a white noise, the curves would all be equal to each other. Term premium would reflect one period of measurable risk premium and nothing else. If the risk premium followed an AR(1), the jumps in each curve would be scaled, converging versions of each other. The same would apply if the risk premium was a random walk (convergence would only be slower). In reality, none of these cases apply at all times. In 2013, the short-term component went negative, and all the other ones were close to zero. During the 2015-16 recession, long-term risk premium was positive and large. In the 2019-20 period, risk premium went negative but only in the first two years; it stabilized thereafter. There is no single pattern.

So how does expected return move in order to generate the patterns estimated by the model? Figure 15 plots the impulse response function of the return factor again (the blue curve). It corresponds to the data-generating probability measure. This time, however, I plot the same-period response of forward rates according to maturity (black curve) and expected interest[13] (red curve). The difference between the two curves corresponds to the term premium. Decomposition (3) implies that such difference depends on risk premia; by construction, risk premia depends on the return factor. Hence, the term premium - the difference between the black and red curves - depends on whether the return factor - the blue curve - is above or below zero.

The figure shows that the non-trivial term structure of risk premia is due to different responses of the return factor to different shocks. For example, in response to a return shock, risk premium increases in the first two years and goes below its mean in the following three years. As a consequence, the term premium should increase in the first two years and decline after that. The plot indicates that this is precisely what happens. Forward rates stay constant - which is consistent with the loadings depicted by figure 9 -, but projected interest rates decline for two year and increase in the next three back to its average. The responses to the other shocks are completely different. We can thus understand how come the model generates a non-trivial term structure of risk premia.

We can also infer from figure 15 that shocks to the level and curve factors are the main drives of

[13]The same-period response of expected interest is the same as the impulse response function of the interest rate itself.



Notes: Each panel plots the impulse response functions of the return factor (blue curve) to the shocks indicated by their title under the data-generating probability measure. The black and red curve indicate the response of the forward rate and the expected interest, respectivell, in period 1, that is, the period in which the impulse hits.

Figure 15: Impulse Response Function of Forward Rate and Interest Rate

long-term premia, but they do not forecast risk premium in the short run. Cochrane and Piazzesi (2009) find similar properties for the slope and curve factors in the US case. Shocks to the return factor do the opposite; they forecast high returns in the short run, but not in the long run.

Unlike the other shocks, innovations to the slope factor have a reduced impact on future expected returns and do not lead to term premia. This is a clarifying result. It explains how return predictability is consistent with the apparent support of the expectations hypothesis provided by Fama-Bliss regressions in section 3. Such support was only apparent. Our coefficient estimates (tables 4 and 5) suggest that forward spreads $f_t^n - i_t$ forecast interest growth, not future returns. The impulse response function of figure 15 now explains how come. Most variation in forward spreads derives from variation in the slope factor, but variation in the slope factor does *not* lead to variation in the return factor. At least, not enough compared to the other three shocks. When running OLS on Fama-Bliss regressions, results reflected this singular trait of the dynamics of returns and prices in the Brazilian market. They did not deny return predictability and they did not prove the validity of the expectations hypothesis.

7 Conclusion

In this article, I use ten years of term structure price data from the Brazilian nominal public bond market. Never before have I had such a long series of data including long-term yields. Our goal, therefore, was to follow some of the research steps already taken for developed markets, in particular the American market.

Our analysis lead to several results, which I now summarize. I find that Fama-Bliss regressions fail to reject the expectations hypothesis, but, in spite of that, returns are highly predictable. After running regressions of returns on past returns and forward rates, I find that the latter have a considerable forecating power. The R^2 reaches 48% in some regressions.

From the analysis of predictability, I compute a single return-forecasting factor for the Brazilian economy. Our factor follows from an intuitive, one-parameter formula that connects short-term interest rates, the short slope of the term structure and its long slope. The return factor captures almost all forecasting power from the traditional level, slope and curve factors calculated from the eigenvalue decomposition of the yield curve.

Based on these findings, I then proceed to estimate an exponential-affine model of the term structure. I impose restrictions on parameters governing the market price of risk that reflect the following assumptions: 1. only the return factor changes risk premium, and 2. only shocks to the level factor are priced. Our estimation focuses on the risk-neutral dynamics of the factors, and infers dynamics under the data-generating probability measure using the market price of risk.

I find that shocks to factors are highly persistent under both probability measures, and that the market price of level shocks is negative. Bonds, therefore, are risky investments. I also characterize factors dynamics by focusing on the co-movements of the return and level factors. The analysis of such co-movement leads us to conclude that, in pricing assets, investors attribute large value to states of nature that lead to prolonged periods of high marginal utility, such as the 2015-16 recession.

I use the estimated model to decompose yields into term premia and expected interest, and the term premia into its different risk-premium components. As a consequence of factors' persistence, I estimate expected interest to be almost as volatile as forward rates, a conclusion that I do not find when estimating dynamics *via* OLS. Finally, the different responses of the return factor to different shocks lead to a complex term structure of risk premia. Based on it, I argue that there is no single term premium, but rather a schedule of term premia according to bond maturity.

I take the present article to be part of initial efforts in the direction of reproducing for the Brazilian market the conclusions - being them the same or not - that were already reached for developed markets. Our results focus on predictability, and highlight the types of macroeconomic shocks that affect marginal utility and bond pricing the most. Tracking their validity over time is an obvious task. As it is the search for structural factors and, hopefully, theoretical models that justify their use on a rigorous basis. The understanding of fixed-income markets is key to the understanding of asset pricing in general and, hence, to macroeconomic dynamics. I hope to contribute to the development of future research in the area, to which alternative questions will not be lacking. There is plenty to be done.

References

T. Adrian, R. K. Crump, and E. Moench. Pricing the term structure with linear regressions. *Journal of Financial Economics*, 110(1):110–138, Oct. 2013. ISSN 0304-405X. doi: 10.1016/j.jfineco.2013.

04.009. URL https://www.sciencedirect.com/science/article/pii/S0304405X13001335.

M. Aguiar and G. Gopinath. Defaultable debt, interest rates and the current account. *Journal of International Economics*, 69(1):64–83, June 2006. ISSN 0022-1996. doi: 10.1016/j.jinteco.2005.05.005. URL https://www.sciencedirect.com/science/article/pii/S0022199605000644.

M. Aguiar and G. Gopinath. Emerging Market Business Cycles: The Cycle Is the Trend. *Journal of Political Economy*, 115(1):69–102, Feb. 2007. ISSN 0022-3808. doi: 10.1086/511283. URL https://www-journals-uchicago-edu.stanford.idm.oclc.org/doi/abs/10.1086/511283.

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

A. Ang, S. Dong, and M. Piazzesi. No-Arbitrage Taylor Rules. Working Paper 13448, National Bureau of Economic Research, Sept. 2007. URL https://www.nber.org/papers/w13448.

M. D. Bauer, G. D. Rudebusch, and J. C. Wu. Correcting Estimation Bias in Dynamic Term Structure Models. *Journal of Business & Economic Statistics*, 30(3):454–467, July 2012. ISSN 0735-0015, 1537-2707. doi: 10.1080/07350015.2012.693855. URL http://www.tandfonline.com/doi/abs/10.1080/07350015.2012.693855.

S. A. Buser, G. A. Karolyi, and A. B. Sanders. Adjusted Forward Rates as Predictors of Future Spot Rates. SSRN Scholarly Paper ID 40165, Social Science Research Network, Rochester, NY, 1996. URL https://papers.ssrn.com/abstract=40165.

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

J. H. Cochrane and M. Piazzesi. Bond Risk Premia. *American Economic Review*, 95(1):138–160, Mar. 2005. ISSN 0002-8282. doi: 10.1257/0002828053828581. URL http://www.aeaweb.org/articles?id=10.1257/0002828053828581.

D. Duffie and R. Kan. A Yield-Factor Model of Interest Rates. *Mathematical Finance*, 6(4):379–406, 1996. ISSN 1467-9965. doi: 10.1111/j.1467-9965.1996.tb00123.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9965.1996.tb00123.x.

E. Fama. The Behavior of Interest Rates. *The Review of Financial Studies*, 19(2):359–379, 2006.

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

E. F. Fama. Forward rates as predictors of future spot rates. *Journal of Financial Economics*, 3(4):361–377, Oct. 1976. ISSN 0304-405X. doi: 10.1016/0304-405X(76)90027-1. URL https://www.sciencedirect.com/science/article/pii/0304405X76900271.

E. F. Fama. The information in the term structure. *Journal of Financial Economics*, 13(4): 509–528, Dec. 1984. ISSN 0304-405X. doi: 10.1016/0304-405X(84)90013-8. URL https://www.sciencedirect.com/science/article/pii/0304405X84900138.

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

M. J. Hamburger and E. N. Platt. The Expectations Hypothesis and the Efficiency of the Treasury Bill Market. *The Review of Economics and Statistics*, 57(2):190–199, 1975.

S. Joslin, M. Priebsch, and K. J. Singleton. Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks: Term Structure Models with Unspanned Macro Risks. *The Journal of Finance*, 69(3):1197–1233, June 2014. ISSN 00221082. doi: 10.1111/jofi.12131. URL https://onlinelibrary.wiley.com/doi/10.1111/jofi.12131.

R. Litterman and J. Scheinkman. Common factors affecting bond returns. *Journal of fixed income*, 1(1):54–61, 1991.

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

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

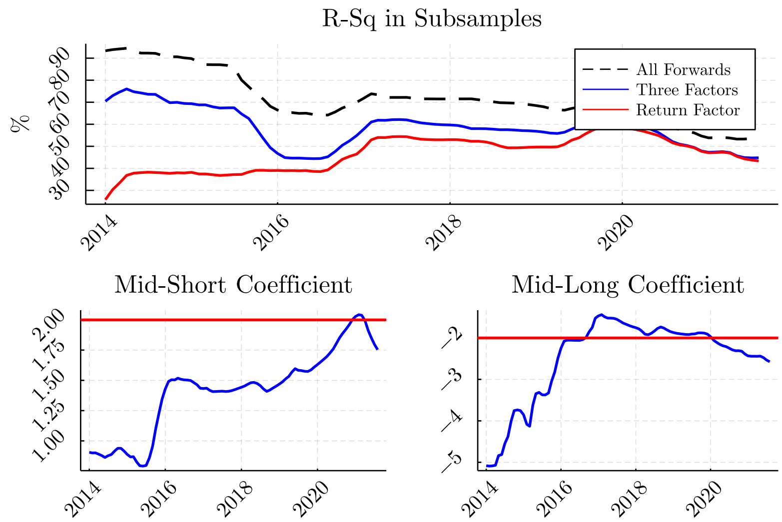
R. J. Shiller, J. Y. Campbell, and K. L. Schoenholtz. Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates. *Brookings Papers on Economic Activity*, 1983(1), 1983.

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

B. M. Tabak and S. C. d. Andrade. Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates. *Brazilian Review of Finance*, 1(1):19–43, 2003. URL https://ideas.repec.org/a/brf/journl/v1y2003i1p19-43.html.

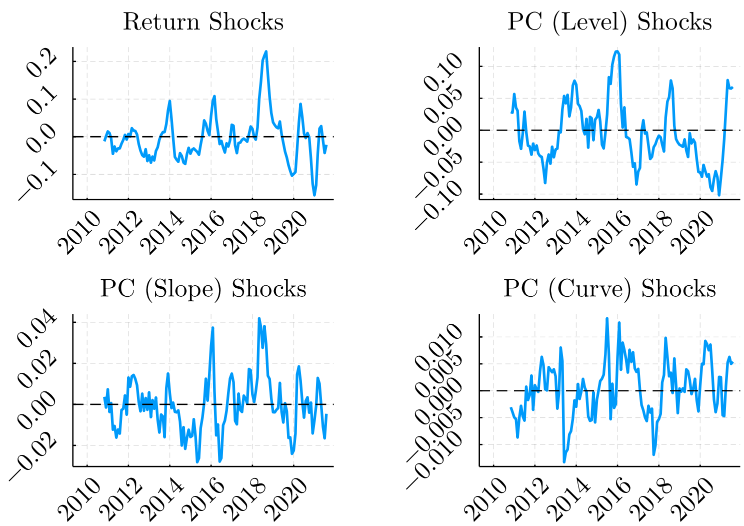
C. Y. Tang and S. X. Chen. Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics*, 149(1):65–81, May 2009.

T. Yamamoto and N. Kunitomo. Asymptotic bias of the least squares estimator for multivariate autoregressive models. *Annals of the Institute of Statistical Mathematics*, 36(3):419–430, Dec. 1984. ISSN 0020-3157, 1572-9052. doi: 10.1007/BF02481980. URL http://link.springer.com/10.1007/BF02481980.



Notes: *Top Panel* I run regressions of the format $\bar{rx}_{t+1} = \alpha + \beta z_t + \epsilon_{t+1}$, where $\bar{rx}_t = \frac{1}{9}\sum_{n=2}^{10} rx_t^n$, in subsamples Sep/2009-T, for various T. In all plots, the horizontal axis indicates the T. The figure plots the regressions' R^2s . I also plot the R^2 for unrestricted regressions on forward rates $\bar{rx}_{t+1} = \alpha + \beta' f_t + \epsilon_{t+1}$. The dashed curve "All forwards" uses all forward rates as independent variables. The "Three Factors" blue curve uses only f^2, f^6 and f^{10} (the same maturities I use to build z - see (5)).*Bottom Panel*: I run regressions $\bar{rx}_{t+1} = \alpha + c_1 f_t^2 + c_2(f_t^6 - f_t^2) + c_2(f_t^{10} - f_t^6) + \epsilon_{t+1}$ for various subsamples Sep/2009-T. The blue curves indicate c_2/c_1 (left panel) and c_3/c_1 (right panel). The red horizontal lines highlight the choice implicit in my baseline build of the return factor: $\theta_s = 2$ and $\theta_l = -2$.

Figure 16: Return Regression in Subsamples 2009-T



Notes: Each panel plots the residuals of OLS regression estimates of the VAR (7).

Figure 17: Estimated Shocks $\hat{e}$ to Factors X

A Additional Tables and Figures

The appendix provides additional results in the form of tables and figures. I make reference to all results in the main text.

Const	**f(1)**	**f(2)**	**f(3)**	**f(4)**	**f(5)**	**f(6)**	**f(7)**	**f(8)**	**f(9)**	**f(10)**	**R-Sq**
-0.0	2.5	-37.1	442.2	-3273.4	14738.0	-41038.5	70637.3	-72737.0	40902.6	-9636.2	33.4
-0.5	*2.1*	*-1.6*	*1.6*	*-1.7*	*1.8*	*-1.9*	*2.0*	*-2.1*	*2.2*	*-2.3*	
-0.0	0.6	0.9	-15.1	33.5	-19.3						28.3
-1.0	*1.2*	*0.2*	*-1.3*	*1.9*	*-2.1*						
-0.0	0.1		-1.5		1.9						7.1
-1.1	*0.5*		*-1.2*		*1.4*						
-0.0		0.0		-3.4		23.9		-31.5		11.3	24.4
-0.5		*0.0*		*-0.6*		*1.4*		*-1.4*		*1.2*	
-0.0		-0.2				4.4				-3.7	22.5
-0.7		*-0.4*				*4.1*				*-3.1*	
-0.0		-0.1								0.3	1.0
-0.3		*-0.1*								*0.4*	
-0.0		-0.5				1.0					7.1
-1.1		*-0.9*				*1.3*					
-0.0						4.3				-3.9	22.3
-0.6						*3.9*				*-3.7*	
0.0		0.1									1.0
0.1		*0.6*									
-0.0						0.3					4.2
-0.6						*1.1*					
-0.0										0.2	1.6
-0.3										*0.8*	

Notes: t statistic in italic. Results from the OLS regression $rx_{t+1}^2 = \alpha + \sum_{n=1}^{10} \beta_n f_t^n + \epsilon_{t+1}$. I use a three month moving average of foward rates on the right-hand side. Each row presents the estimates for a different set of explanatory variables. R^2 adjusted for model parameterization, and standard deviations consistent for error heteroskedasticity.

Table 12: 2y Bond Return Forecast Regression on Forward Rates (3-Month Moving Avg.)

Const	f(1)	f(2)	f(3)	f(4)	f(5)	f(6)	f(7)	f(8)	f(9)	f(10)	R-Sq
-0.8	10.0	-165.0	2040.3	-16148.2	77185.5	-225571.0	402865.6	-426283.2	244498.1	-58424.6	52.8
-2.7	*0.9*	*-0.8*	*0.8*	*-1.0*	*1.1*	*-1.3*	*1.5*	*-1.6*	*1.8*	*-1.9*	
-1.0	1.3	1.8	-71.9	145.9	-68.1						49.9
-3.6	*0.4*	*0.1*	*-0.9*	*1.1*	*-0.9*						
-1.0	-1.0		-15.6		25.5						43.2
-3.9	*-0.7*		*-2.6*		*3.4*						
-0.9		-4.4		-30.2		187.2		-242.9		98.5	50.1
-3.3		*-0.9*		*-1.1*		*1.9*		*-1.8*		*1.6*	
-1.0		-6.9				32.1				-16.3	48.6
-3.4		*-2.1*				*4.2*				*-2.0*	
-0.9		-5.9								13.2	30.9
-3.6		*-1.4*								*2.5*	
-1.1		-8.0				17.0					44.3
-3.9		*-2.4*				*3.5*					
-0.7						29.1				-22.7	37.5
-2.8						*3.0*				*-2.4*	
-0.2		2.6									11.1
-1.5		*1.9*									
-0.7						6.1					28.4
-3.1						*3.5*					
-0.6										5.3	23.0
-2.9										*3.4*	

Notes: t statistic in italic. Results from the OLS regression $rx_{t+1}^{10} = \alpha + \sum_{n=1}^{10} \beta_n f_t^n + \epsilon_{t+1}$. I use a three month moving average of foward rates on the right-hand side. Each row presents the estimates for a different set of explanatory variables. R^2 adjusted for model parameterization, and standard deviations consistent for error heteroskedasticity.

Table 13: 10y Bond Return Forecast Regression on Forward Rates (3-Month Moving Avg.)

Const	**f(1)**	**f(2)**	**f(3)**	**f(4)**	**f(5)**	**f(6)**	**f(7)**	**f(8)**	**f(9)**	**f(10)**	**R-Sq**
-0.0	2.0	-24.6	251.4	-1744.3	7553.0	-20329.6	33862.2	-33776.8	18427.8	-4220.7	23.3
-0.7	*2.9*	*-2.3*	*2.1*	*-2.1*	*2.2*	*-2.2*	*2.3*	*-2.4*	*2.4*	*-2.4*	
-0.0	0.8	-2.3	-1.8	11.3	-7.5						20.4
-1.1	*2.0*	*-0.8*	*-0.2*	*0.9*	*-1.2*						
-0.0	0.0		-0.9		1.3						4.9
-1.0	*0.1*		*-0.9*		*1.2*						
-0.0		-0.5		0.4		7.6		-10.2		3.0	16.0
-0.7		*-0.6*		*0.1*		*0.6*		*-0.7*		*0.5*	
-0.0		-0.3				3.0				-2.3	14.8
-0.7		*-0.6*				*3.3*				*-2.4*	
-0.0		-0.1								0.3	0.8
-0.3		*-0.2*								*0.6*	
-0.0		-0.4				0.8					6.1
-1.0		*-0.9*				*1.4*					
-0.0						2.8				-2.5	14.1
-0.5						*3.1*				*-2.8*	
0.0		0.1									0.5
0.2		*0.5*									
-0.0						0.3					3.5
-0.5						*1.1*					
-0.0										0.2	1.4
-0.2										*0.8*	

Notes: t statistic in italic. Results from the OLS regression $rx^2_{t+1} = \alpha + \sum_{n=1}^{10} \beta_n f^n_t + \epsilon_{t+1}$. Each row presents the estimates for a different set of explanatory variables. R^2 adjusted for model parameterization, and standard deviations consistent for error heteroskedasticity.

Table 14: 2y Bond Return Forecast Regression on Forward Rates (Not Moving Avg)

Const	f(1)	f(2)	f(3)	f(4)	f(5)	f(6)	f(7)	f(8)	f(9)	f(10)	R-Sq
-0.3	5.9	-78.1	816.4	-5843.4	26039.7	-71871.6	122229.9	-123951.9	68501.1	-15845.1	39.9
-2.3	*2.2*	*-1.8*	*1.8*	*-1.8*	*2.0*	*-2.1*	*2.2*	*-2.3*	*2.3*	*-2.4*	
-0.4	1.9	-7.2	-6.2	31.1	-16.2						37.1
-2.9	*1.4*	*-0.7*	*-0.2*	*0.6*	*-0.6*						
-0.3	-0.5		-5.0		8.5						27.9
-2.8	*-0.5*		*-1.9*		*2.8*						
-0.3		-2.8		-2.0		40.6		-51.9		18.9	35.9
-2.5		*-1.1*		*-0.1*		*1.0*		*-1.0*		*0.9*	
-0.3		-2.5				12.4				-6.8	34.5
-2.5		*-1.7*				*3.7*				*-1.9*	
-0.3		-1.8								4.3	17.5
-2.5		*-1.1*								*2.0*	
-0.4		-2.9				6.1					29.5
-2.9		*-2.0*				*3.0*					
-0.2						10.9				-8.6	26.7
-1.8						*3.0*				*-2.4*	
-0.1		0.9									5.9
-0.8		*1.4*									
-0.2						2.2					18.2
-2.0						*2.5*					
-0.2										1.9	13.5
-1.8										*2.3*	

Notes: t statistic in italic. Results from the OLS regression $rx^5_{t+1} = \alpha + \sum_{n=1}^{10} \beta_n f^n_t + \epsilon_{t+1}$. Each row presents the estimates for a different set of explanatory variables. R^2 adjusted for model parameterization, and standard deviations consistent for error heteroskedasticity.

Table 15: 5y Bond Return Forecast Regression on Forward Rates (Not Moving Avg)

Const	**f(1)**	**f(2)**	**f(3)**	**f(4)**	**f(5)**	**f(6)**	**f(7)**	**f(8)**	**f(9)**	**f(10)**	**R-Sq**
-0.9	9.8	-137.0	1434.9	-10457.9	47483.7	-133095.9	228850.4	-233759.3	129778.4	-30099.4	47.8
-3.3	*1.7*	*-1.5*	*1.4*	*-1.5*	*1.7*	*-1.8*	*1.9*	*-2.0*	*2.1*	*-2.1*	
-1.0	3.2	-18.3	13.9	-0.7	10.5						45.8
-3.9	*1.1*	*-0.8*	*0.2*	*-0.0*	*0.2*						
-0.9	-1.1		-12.8		22.0						42.4
-4.0	*-0.7*		*-2.6*		*3.8*						
-0.9		-6.3		-8.0		77.0		-90.5		35.7	45.2
-3.7		*-1.2*		*-0.3*		*0.9*		*-0.9*		*0.9*	
-0.9		-6.6				22.7				-7.9	45.0
-3.7		*-2.3*				*3.4*				*-1.2*	
-0.8		-5.4								12.4	32.6
-4.0		*-1.7*								*3.0*	
-1.0		-7.1				15.4					43.9
-4.0		*-2.5*				*3.8*					
-0.6						18.6				-12.6	32.5
-2.8						*2.4*				*-1.7*	
-0.2		2.5									10.5
-1.5		*1.9*									
-0.6						5.8					28.7
-3.0						*3.5*					
-0.6										5.2	24.1
-2.9										*3.4*	

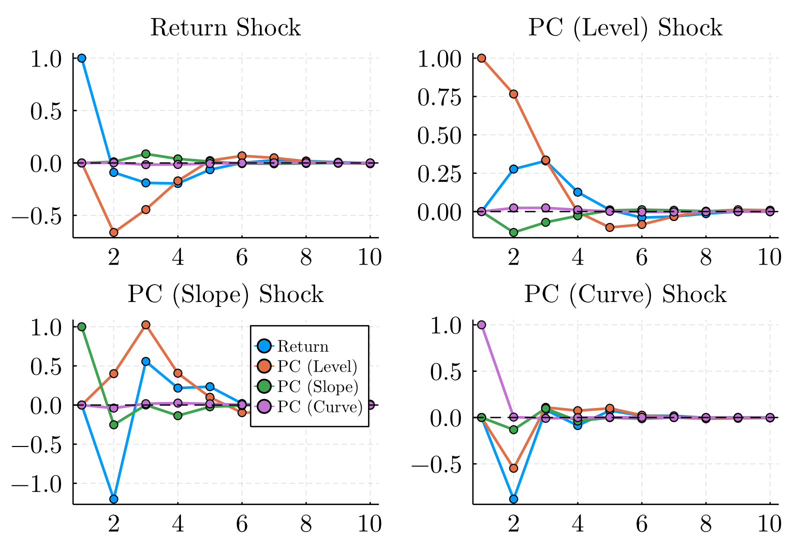
Notes: t statistic in italic. Results from the OLS regression $rx_{t+1}^{10} = \alpha + \sum_{n=1}^{10} \beta_n f_t^n + \epsilon_{t+1}$. Each row presents the estimates for a different set of explanatory variables. R^2 adjusted for model parameterization, and standard deviations consistent for error heteroskedasticity.

Table 16: 10y Bond Return Forecast Regression on Forward Rates (Not Moving Avg)

Const	z	PC 1	PC 2	PC 3	PC 4	PC 5	R-Sq
0.0	0.2						15.5
2.5	*2.2*						
0.0	0.2	-0.0					14.9
2.5	*2.0*	*-0.1*					
0.0	0.3	-0.0	-0.5				19.3
2.5	*3.3*	*-0.5*	*-1.5*				
0.0	0.3	-0.1	-0.5	-0.4			19.5
2.5	*3.5*	*-0.6*	*-1.6*	*-0.9*			
0.0		0.1					2.0
2.4		*0.8*					
0.0		0.1	0.2				2.4
2.5		*0.8*	*0.5*				
0.0		0.1	0.2	0.1			1.7
2.5		*0.8*	*0.5*	*0.2*			
0.0		0.1	0.1	0.2	2.8		20.0
2.5		*1.2*	*0.2*	*0.5*	*4.2*		
0.0		0.1	0.1	0.2	2.8	3.1	22.2
2.4		*1.2*	*0.2*	*0.5*	*4.3*	*1.3*	

Notes: Results from the regression $rx^2_{t+1} = \alpha + \beta z_t + \delta' pc_t + \epsilon_{t+1}$, where pc_t groups the term structure principal component (see subsection 2.4). Each row corresponds to a different specification. t statistic in italic. R^2 adjusted for model parameterization, and standard deviations consistent for error heteroskedasticity.

Table 17: 2y-Bond Return Forecast Regression on Return Factor and Principal Components



Notes: Each panel plots the impulse response function of the model's factors to the shock indicated by the corresponding title.

Figure 18: Impulse Response Functions to factors' VAR - OLS Estimation

Const	z	PC 1	PC 2	PC 3	PC 4	PC 5	R-Sq
0.0	2.2						47.8
1.2	*4.9*						
0.0	2.0	0.5					49.3
1.1	*4.0*	*0.9*					
0.0	1.8	0.5	0.5				49.0
1.2	*4.1*	*1.0*	*0.3*				
0.0	1.8	0.5	0.6	1.4			48.8
1.2	*3.6*	*1.0*	*0.3*	*0.5*			
0.0		1.3					18.5
1.0		*2.7*					
0.1		1.2	5.0				35.8
1.4		*2.5*	*2.4*				
0.1		1.3	4.8	4.8			37.1
1.3		*2.6*	*2.4*	*1.5*			
0.0		1.5	4.3	5.5	15.3		45.5
1.2		*2.9*	*2.5*	*2.2*	*3.0*		
0.0		1.5	4.3	5.5	15.4	10.2	45.5
1.1		*3.0*	*2.5*	*2.2*	*3.0*	*0.6*	

Notes: Results from the regression $rx^{10}_{t+1} = \alpha + \beta z_t + \delta' pc_t + \epsilon_{t+1}$, where pc_t groups the term structure principal component (see subsection 2.4). Each row corresponds to a different specification. t statistic in italic. R^2 adjusted for model parameterization, and standard deviations consistent for error heteroskedasticity.

Table 18: 10y-Bond Return Forecast Regression on Return Factor and Principal Components

	Const.	Return	PC (Level)	PC (Slope)	PC (Curve)	|Eig|	R-Sq
Return	0.00	-0.09	0.28	-1.20	-0.88	0.57	0.19
PC (Level)	0.00	-0.66	0.77	0.40	-0.55	0.57	0.47
PC (Slope)	0.00	0.01	-0.14	-0.25	-0.13	0.40	0.34
PC (Curve)	0.00	0.00	0.02	-0.04	0.00	0.03	0.09

Notes: I estimate the VAR model (7) by OLS. The table reports my estimates for μ, Φ, the absolute value of eigenvalues of Φ and the R^2 of each regression.

Table 19: Parameter of the factors' VAR - OLS Estimation