Abstract

This report studies the predictive power of the empirical slope of the structure of interest rates, that is the difference between 10-years and 3-months bonds, when forecasting macroeconomic variables in emerging countries, and tries thus to answer *"Forecasting macroeconomic variables through the term-structure of interest rates in emerging countries: how did the Great Recession affect the term spread predictive power across the globe?"*

Indeed, the literature having not been able to settle on a consensus, still describing the term spread as a *"stylized fact"*, besides focusing mostly on the United Stated and members of the European Union, this paper tries to fill in the gap by comparing multiple emerging countries (11) forecast performances for month-over-month unemployment growth rate, inflation growth rate and the industrial production growth rate.

Moreover, the period selected, from 2000-2005 to 2016-2018, should result in structural breaks which in turn should improve the said predictive ability, as suggested by Wheelock and Wohar (2009, p.430) in their survey of past results, highlighting the loss of it ever since the mid-1980s and an increased stability in macroeconomic factors.

However, if this paper finds evidence of differences across countries, even across regions (as underlined by Asian countries results when considering the inflation growth), but also evidence supporting the *"intertemporal consumption theories"* (through the unemployment rate for which Russia only, due to its economical and political system differing from the others, appears to bear some significativity, at least in-sample), or even of the importance of the monetary regime (strict control of inflation worsening performances as illustrated by the Russian and South African cases pseudo out-of-sample), it was unable to confirm the thesis.

Indeed, as demonstrated by South Africa, for which data were voluntary chosen to start in 1981 instead of after the beginning of the second millennium, the structural break of the mid-1980s, reflecting new monetary policy stance, still proves itself to be more significant, highlighting only the little evidence supporting the idea of improvement in predictive power of the term spread following the Great Recession.

Table of content

1. Introduction

The literature not being able to reach a consensus towards best practices for the term spread when forecasting macroeconomic indicators, as well highlighting contradictory results, going as far as describing it as a *"stylized fact"*, this paper takes another, different look (by cross-checking results across the globe) at the forecasting potential of the term-structure of interest rates.

Furthermore, as most authors agree, outcomes can vary from model to model, but more specifically from country to country, and through time. Additionally, they mostly focused their researches on the United Stated and some nations member of the European Union (mostly highly advanced ones such as Germany or France).

At last, the Great Recession, most likely having impacted macroeconomic factors stability, the structural break (described further in this report), taking place in the mid-1980s, era when such factors started display an increased stability and leading to worsened forecast performances, could be of the past due to renewed volatility, hence improving the term spread predictive ability, as suggested by past authors such as Wheelock and Wohar (2009, p.430).

At the light of those three elements, and especially the latter, it is quite naturally that the following question imposed itself:

"Forecasting macroeconomic variables through the term-structure of interest rates in emerging economies: how did the Great Recession affect the term spread predictive power across the globe?"

Accordingly, the thesis developed to answer that question is none other than the one explained here above, being the renewed instability in macroeconomic indicators due to the Great Recession, leading to renewed significant predictive power for the term spread.

Therefore, the three hypothesis tested are the following:

- A significant structural break taking place during or around the Great recession.
- As underlined by researchers, results vary from country to country, hence are not expected to be standardized but to differ based on the nation's political system, economy, etc.

- Given the importance of monetary-policy, countries displaying a strict control over the inflation growth rate should have worse performances than the others.

To answer those, the paper will first go through some literature review, providing key elements to understand and interpret the results, before presenting the data and quickly analysing the countries composing the database, followed by the methodology, and finalised by the analysis of the model performances.

Finally, in order to build the report, as it will be explained later on, I first had to collect the data, adjust them, build the models and collect the performances.

Illustrating those two quotes:

"All models are wrong. Some are useful."

George Box.

"The only function of economic forecasting is to make astrology look respectable."

John Kenneth Galbraith.

Difficulties were encountered at every step along the way, being the program itself, as it happens with eViews, or when building the models as they can always be updated in order to try to improve them.

As explained later, this paper finally settled on a quite simple univariate linear model, applicable for monthly data to all countries while respecting stationarity test, in order to grasp the term spread predictive ability in its mere form.

2. Literature review

2.1. Why consider the yield curve?

The structure of interest rates, also called in the literature yield curve, and more specifically its slope, the term spread, studied in this paper, has long been seen as one of the most consistent predictors of future economic activity despite no universal consensus between the researchers. Indeed, results and related conclusions vary from author to author and more specifically, from country to country and period considered to period considered. As such, Benati and Goodhart (2008, p.1237) described this trend of forecasting macroeconomic variables through the yield curve as a *"stylized fact in search of a theory."*

According to Investopedia, the term structure of interest rates can be defined as such:

The relationship between interest rates or bond yields and different terms or maturities. When graphed, the term structure of interest rates is known as a yield curve, and it plays a central role in an economy. The term structure reflects expectations of market participants about future changes in interest rates and their assessment of monetary policy conditions. (Chen, 2009)

Quite logically, one would tend to assume, as explained, among others, by Wheelock and Wohar (2009, p.423) that the expectations hypothesis underlying the term spread is the most obvious reason for considering the said tool while forecasting macroeconomic variables. That is, long-term yields consist in summing future short-term rates (up to the period considered) plus a term premium, main reason behind the upward behavior of the slope. However, the slope will flatten or invert if public expectations anticipate falling yields.

As described by many authors and mentioned in Investopedia's definition as well, short-term interest rates reflect monetary policies, whereas long-term ones hold public expectations. Bauer and Mertens (2018, p.2) explained that *"while long-term move up with short-term during the early part of an expansion, they tend to stop doing so once investors' economic outlook becomes increasingly pessimistic."* If this seems to be part of the story to explain the use of this tool, literature doesn't stop there.

Indeed, as highlighted by Wheelock and Wohar (2009, pp.423-424) in their survey of the current literature on the topic, two main theories exist to explain this phenomenon: the *"monetary-policy"* and the *"theories of intertemporal consumption"*.

The former tends to consider adjustments as results of the said monetary policies, for instance presenting a tightening as a possible cause of curve inversion/flattening. Nevertheless, as surveyed, Feroli (2004), Estrella (2005), Estrella and Trubin (2006) did emphasise the importance of the *"monetary authority's policy objectives and reaction function"* in determining the predictive ability of the spread. Its choice of controlling either the macroeconomic variable or inflation hence impact related results. Moreover, variation in the authority responsiveness could further affect them, improving or worsening their forecasting ability, depending on the change.

The latter *"derives a relationship between the slope of the yield curve and future economic activity explicitly from the structure of the economy"* (Wheelock and Wohar, 2009, p.423). To defend it, Harvey (1988) assumed that individuals prefer stability over volatile consumption and would rather spread out their revenues to ensure constant income and goods of consumption, and not just only in period of growth. As far as this theory is concerned, however, much of the empirical evidence about the structure of interest rates is tied to the nominal, hence requiring persistence of inflation for the forecasting results to be consistent.

While the literature still tends to describe the yield curve as a "*stylized fact in search of a theory*", most authors agreed upon its usefulness. Indeed, while many have found a decrease in its predictive ability since the mid-1980s (among which D'Agostino, Giannone and Surico (2006), or Haubrich and Dombrosky (1996)) attributed to a greater stability in the economy, Estrella (2005, p.27) still found that though not structural, predictive relationships are robust. He concludes that *"the yield curve has predictive power, for example, if the monetary authority follows "strict" or "flexible" inflation targeting, as defined by Svensson (1999)."* Nonetheless, based on evidence of the term's spread significative predictive power independently of other measures, Benati and Goodhart (2008) did suggest that the monetary-policy theory alone couldn't explain such results.

2.2. Past results

The first element to point out here again regarding the results, alike the reason for considering the yield curve when forecasting macroeconomic variables, is that no consensus has been reached among the researchers despite a general agreement regarding its usefulness.

Undeniably, as emphasised by Wheelock and Wohar (2009) in the case of output growth at the light of other researchers' results, and confirmed by Dewachter, Iania and Lyrio (2014, p.10) *"the results in the literature regarding the importance of each yield spread component are contradictory."* This does include the empirical slope considered in this paper[1].

Reasons most often advanced range from the type of model, to the timeframe considered, as well as the country studied, highlighting the idea of *"stylized fact"* exposed by Benati and Goodhart (2008).

Due to the importance of monetary policies, the latter comes as no surprise. However, despite what could be suspected with standardised results across similar countries, the literature does show again contradictions. Indeed, for instance, Davis and Fagan (1997) did find statistical evidence for the term spread when studying the output growth in Europe, contradictory to a related research led by Berk and van Bergeijk (2001), suggesting that little information was contained by the spread. The same contrast was observed between Hu's conclusion for Japan (1993) and Harvey's (1991).

Aside from model differences most likely explaining those inconsistencies, Shareef and Shijin (2017), while studying the yield curve and macroeconomic factors in India, found evidence of influence by the fiscal deficit on short-term rates and by yield expectations review on long-term rates. It thus seems safe indeed to assume that monetary-policy alone do not explain all, but that other factors reflecting the economic (in)stability, which affects the curve predictive power, should be considered as well.

Moreover, components considered can vary from one researcher to another. Differences can be found regarding the spread only. Some, as this paper does, looked at the empirical slope (Hännikäinen, 2016) while others considered each separate component, hence short and long-term rates, besides adding control variables (Dewachter, Iania, Lyrio, 2014; Estrella, 2005; …) If most authors agree upon the significance of the spread distinctive components, they also find

[1] Difference between the 10-years and 3-months bonds

adding control variables of little use. Others, however, found evidence of the term spread's significance despite adding short-term rates[2].

Furthermore, as no forecasting technique has been universally accepted, models differ as well from paper to paper. From univariate linear models, to multivariate, to non-linear ones such as neural network, much was imagined, bringing new contributions to the literature. This paper will focus on a univariate linear model. Though, others have served to demonstrate structural breaks in the data as well as nonlinearities, lack of usefulness when adding control variables, etc.

As an example, in their survey of the literature, Wheelock and Wohar (2009) found that if most agreed on the significance of the spread when forecasting the output growth 6 to 12-months ahead, differences are observed between the authors' surveyed conclusions. Regarding the empirical slope only, Dewachter, Iania and Lyrio (2014, p.12) found that *"the yield spreads alone are statistically significant only for a horizon of 8 quarters with minor predictive ability."*

At last, frequency of the data measured could also explain the differences, as some use quarterly whereas others prefer working with monthly.

Let us consider now the importance of the timeframe of the set. Indeed, multiple researchers (Dotsey, 1998; Estrella, Rodrigues and Schich, 2003) have emphasised the break occurring in the mid-1980s and reducing the term spread predictive ability.

The common explanation lays within the idea that since that period, up until 2007, the improved steadiness of macroeconomic variables have impacted the forecast results, reason why this paper explores the empirical slope performance based on data corresponding to that period, as it will be further developed later.

To confirm this impression, Bordo and Haubrich (2004, 2008) discovered a positive correlation between inflation persistence and the forecast accuracy, in the meantime finding several breaks in the considered relationship, associated with changes in the monetary regime or correlated to the inflation persistence.

[2] Reflecting monetary policies

3. A look at the data

3.1. Data description

The dataset considered in this paper is composed of 11 emerging countries (as defined by LSM (2018), except for Israel, Kenya and Uruguay) for which the term spread and the inflation were retrieved, as well as the unemployment rate and the industrial production level for those for whom those data were available (leading to a total of 6.077 observations).

Indeed, this report relying on free data only, access was not so easy, explaining the lack of common further variables as well as timeframe. The data sources were both Investing.com for most bond data, as well as OECD.com for remaining ones besides the macroeconomic indicators studied here.

As mentioned, the timeframe here studied was voluntary chosen as to englobe the recession of 2007 and to analyse its impact on the slope predictive power through the following decade. As such, data do start in between 2000 and 2005 for most countries, ending in between 2016 and 2018, with the exception of South Africa, starting in 1981, allowing for a sanity check regarding the idea of a break in the forecasts since the mid-80s.

Accordingly to most works read through, the term spread is built on zero-coupon bonds of maturity of 10 years and 3 months. As for the other variables (industrial production level, inflation, unemployment rate), the frequency chosen was monthly.

Not studied here, authors agreed on recessions being usually well forecasted by the term spread. In fact, no later than last May, the New-York Times (NYT, 2019) was releasing an article about the yield curves dangerously inverting and that investors should keep an attentive eye on the situation.

Indeed, as for the others, a clear relationship can be deduced between the two quoted, recessions occurring after every curve inversion, as highlighted by the NYT. Logically, one would imagine inflation, industrial production and the spread to be positively correlated, whereas a negative one is expected with the unemployment rate, usually rising during recessions. Nevertheless, no statistical evidence could be found to support this deduction.

To summarize, the database is composed of the following countries, containing the following variables and with the related number of observations:

- **Brazil**: from 01/01/2006 to 01/09/2017 (141 observations per variable); all 4 variables.
- **China**: from 01/01/2003 to 01/05/2015 (149 obs./var); 2 variables (spread & inflation).
- **India**: from 01/02/2002 to 01/11/2017 (190 obs./var); 3 variables (spread, industrial production & inflation).
- **Indonesia**: from 01/08/2004 to 01/06/2018 (167 obs./var); 2 variables (spread & inflation).
- **Israel**: from 01/01/2001 to 01/03/2018 (207 obs./var); 3 variables (spread, industrial production & inflation).
- **Kenya**: from 01/03/2006 to 01/08/2018 (150 obs./var); 2 variables (spread & inflation).
- **Mexico**: from 01/01/2002 to 01/09/2017 (189 obs./var. except for unemployment rate (96 obs.), stopping on 01/12/2009); all 4 variables.
- **Russia**: from 01/01/2002 to 01/04/2018 (196 obs./var. except for unemployment rate (135 obs.), starting on 01/02/2007); all 4 variables.
- **South Africa**: from 01/01/1981 to 01/07/2018 (451 obs./var.); 2 variables (spread & inflation).
- **South Korea**: from 01/01/2001 to 01/07/2018 (211 obs./var.); all 4 variables.
- **Uruguay**: from 01/02/2004 to 01/12/2016 (155 obs./var.); 3 variables (spread, inflation & unemployment rate).

3.2. Country analysis

As highlighted before, monetary policies do not bear the full weight of significance for the structure of interest rates forecast results, and although summarized, it is interesting to have a look at the countries' respective economies, politics, etc.

First, a distinction can obviously be made between China, Russia and the others. Indeed, due to their political system, communist, those two operate in economies that differ from the rest, working in free market economies. As for the two quoted, China has the use of a social market economy whereas Russia, while undergoing transformation towards a free market, still uses a centrally planned economy. It will be interesting to analyse those two results, at the light

of the others', especially regarding unemployment rate as the system should influence its behavior.

Second, as explained by Shareef and Shijin (2017, p.1), the concept of fiscal deficit could be of importance. Indeed, the authors suggested that *"short-term interest rates are mainly influenced by the fiscal deficit present in emerging economies while long-term rates get affected when market participants revise their expectations on yields."* Although the structure used for the model of this report is not decomposed, these could still be potential reasons for differences between the nations' respective results.

Nevertheless, through TradingEconomics.com, we can notice that there are no significant variation in terms of fiscal deficit, all ranging in between -1.20% and -4.20%, apart from Brazil and Kenya a bit lower at -7.10% and -6.70% respectively, and Russia, only country with a positive balance, for the second time over the past several years (only one in this case). Therefore, fiscal deficit shouldn't be a major source of differentiation.

As for the public debt, significant changes can be observed from one location to another. Truly, when Brazil peaks at 77%, Russia bottoms at around 13.5%, only country under 25%. For the rest, Indonesia, South Korea and Mexico manage to remain within the 25-50% bracket, whereas the others are all included within the 50-75% except thus for Brazil.

Finally, if past monetary policies will not be fully analysed here, general trends can still be observed thanks to the data. Indeed, as displayed on *Figure 1*, it would appear that the month-over-month unemployment rate growth does show a high volatility, which is confirmed by data for other countries. As far as the month-over-month industrial production growth as well as inflation's are concerned, trends vary from one country to another, most likely due to the government monetary policy, strictly controlling inflation or not. Based on graphical displays, one could assume Israel, Russia, South Africa and Uruguay to keep a close eye on their inflation levels, which should worsen their outcomes[3].

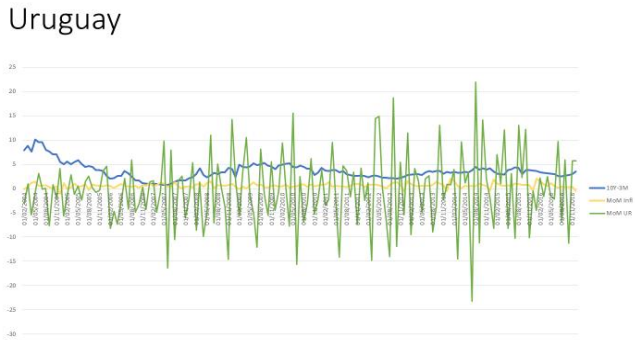

Figure 1: Evolution of the spread, the monthly growth of inflation and unemployment rate in Uruguay.
(Green= Unemployment rate; Blue= Spread; Yellow= Inflation)

[3] Confirmed by the country respective central bank website

4. Methodology

4.1. Manipulations

Once the set complete, first manipulations occurred. First, data underwent a STL decomposition[4] as defined by Cleveland, Cleveland, McRae and Terpenning (1990), splitting series into its trend, seasonal and remainder components. The following was executed in eViews, from which were extracted the seasonally adjusted data composing the current set.

Furthermore, after having adjusted the set, if, as mentioned previously, the term spread was built upon the monthly difference of the two distinctive bonds[5], growth rates were defined for the other variables thanks to the following formula:

$$\Delta Y_t = \log(\frac{Y_t}{Y_{t-1}})$$

Unlike Wheelock and Wohar, rates, in the case of this paper, were computed on a month-over-month basis instead of a year-over-year. If doing so most likely increased the volatility of the time series selected, as observed on *Figure 1*, it does ensure nonetheless their robustness when undergoing the Augmented Dickey-Fuller test (ADF test), unlike for year-over-year rates, regardless of the horizon considered. As explained by Arnold, Gerber, Hanck and Schmelzer (2018, p. 335), the said test checks for unit root when auto-regressing a variable, meaning that the series are stationary or not, required to avoid misleading inferences.

The same test was applied to the models to ensure the absence of unit root in the residuals. Again, if month-over-month growth produced robust results across all horizons, it was not the case for year-over-year ones, at the very least for the model considered here.

4.2. Model

Although nonlinear models appeared to find more and more users ever since the discovery of a break in data in the mid-80s, the model selected for this paper falls down into the linear category, much like the one summarized by Wheelock and Wohar (2009, p.425):

$$\Delta Y_{t+k} = \alpha + \beta\, Spread_t + \varepsilon_{t+k}$$

[4] Seasonal and Trend decomposition using Loess
[5] 10-years and 3-months

Where k is the horizon forecasted and is included between {1;18} (up to 6 quarters ahead thus); ΔY_{t+k} is the growth rate of the macroeconomic variable considered at the horizon k; α is the intercept; $Spread_t$ the defined difference between the 10-year and 3-month bonds; and finally, ε_{t+k}, Newey-West Errors at the horizon k.

In their book, Arnold and al (2018, pp. 357-358) highlighted the usefulness of those *"Heteroskedasticity-and autocorrelation-consistent (HAC) estimators of the variance-covariance matrix."* Indeed, Newey-West Errors allow to account for autocorrelation as well as heteroskedasticity in the data, ensuring robustness of the results. The authors further define those estimators as the following:

$$\tilde{\sigma}^2_{\widehat{\beta_1}} = \hat{\sigma}^2_{\widehat{\beta_1}} \hat{f}_t$$

Where $\hat{\sigma}^2_{\widehat{\beta_1}}$ *is the heteroskedasticity-robust variance estimate of* $\widehat{\beta_1}$ *and*

$$\widehat{f_t} = 1 + 2 \sum_{j=1}^{m-1} \left(\frac{m-j}{m}\right) \tilde{\rho}_j$$

Is a correction factor that adjust for serially correlated errors and involves estimates of m-1 autocorrelation coefficients $\tilde{\rho}_j$*. For a time series X, we have:*

$$\breve{\rho}_j = \frac{\left(\sum_{t=j+1}^{T} \hat{v}_t \hat{v}_{t-j}\right)}{\sum_{t=1}^{T} \hat{v}_t^2}, \text{ with } \hat{v} = (X_t - \bar{X})\hat{u}_t.$$

Finally, m is a truncation parameter to be chosen. A rule of thumb for choosing m is

$$m = 0.75 * T^{\frac{1}{3}}, \text{ with T the number of observations.}$$

This paper looking firstly for evidence of predictive power in the empirical slope despite the era analysed, tables summarizing results for Granger causality tests can be found in the appendices. This concept, based on prediction, helps determine if a series is useful when forecasting another one. Henceforth, one would be said to "G-cause"/precede another one, if the former comes before the latter as well as demonstrates a causal link/relationship with it.

Furthermore, as most authors highlighted improvements in their results when applying out-of-sample forecasts, all models[6] have been Granger tested in-sample and Single Sample t-tested pseudo out-of-sample. Kent University (2009, para Common uses) supported the use of it when testing "*statistical difference between a change score and zero."* The change score consists of the difference between the out-of-sample observation and the forecasted result. Once built, its mean is then compared to zero in order to check for significative change in the series, meaning if the forecast is right or wrong on average, by looking at *"potential changes between two time points for the original measures."*

It is important to single out that as a rule of thumb, all models were required to include at the very least 60 observations per existing variable and pseudo out-of-sample forecasts to contain 40 observations (unless exception).

4.3. Accounting for a break in data

Explained earlier, evidences of structural breaks and nonlinearities have been flourishing in the literature since the 1990s. As such, if this research did not look further for nonlinearities, it did account for a break. Indeed, Stock and Watson (1996) demonstrated that model instability diminished its performances. Given the period considered, from 2000-2005 to 2016-2018, instability can be expected.

Besides, as noted earlier, a change in monetary policy, potentially resulting from the Great Recession[7], could impact the predictive ability of the term spread. The South African case, starting in 1981, will allow a comparison between the discussed break of the mid-1980s and a potential break taking place in the 2000s to confirm the hypothesis of improved forecast performances.

As break dates were unknown, a Quandt Likelihood Ratio (QLR) test was used. The break is determined by retrieving the one date with the maximum F-statistics obtained when applying the test.

[6] Understand all parametric variations of the model presented. Variations involve the horizon, the sample length, as well as the macroeconomic variable considered.

[7] A common agreement regarding its period has it last from 2008 to 2012.

Here again, all models have been forecasted both in-and out-of-sample, with the respective results retrievable in the appendices.

5. Results presentation

Before looking deeper at the results presented in tables 1 to 18 within the appendices, it is important for the reader to understand the hypothesis behind the tests leading to the discussed outcomes.

If, for the Granger causality test (which, due to the simplified model used, provides results equalling models' spread p-value), the objective is to reject the null hypothesis, being that the empirical slope has no power explaining the considered growth rate, the Single Sample t-test requires the inability of the model to reject the null hypothesis, being that the true mean of the change score equals zero, to support the thesis of better performances resulting from a greater instability in macroeconomic variables following the Great Recession.

5.1. Industrial production growth rate

Forecast performances of the term spread when looking at industrial production growth rate provide little evidence to confirm the idea behind the paper, regardless of the horizon considered[8].

Indeed, regarding in-sample performances for the full set[9] contained in *Table 1*, out of 108 results, five only appeared to be significant at a 5% level, that is for which the null hypothesis of the Granger causality test is rejected:

- Three in South Korea: {t+1; t+2; t+3}
- Two in India: {t+1; t+4}

If taking into account a larger value (10% as a threshold), 8 more results could be considered as significant: one in Brazil, {t+3}; another in India, {t+3}; four in Mexico, {t+1; t+2; t+3; t+5}; one in Russia, {t+1}; and a last one in South Korea, {t+4}.

[8] The countries in the set are the following: Brazil, India, Israel, Mexico, Russia, South Korea
[9] No window applied, that is no break accounted for.

As for the Standard Error (also called SER), relatively stable performances can be observed both for results concerning one country only as well as across the nations' respective outcomes, ranging overall from 0.016 to 0.248.

Pseudo out-of-sample performances in *Table 2* would, overall, tend to support the assumption of a true mean equals zero. Indeed, if Mexico's Single Sample t-test results strongly reject the null hypothesis for all 18 periods considered[10], apart from Brazil {t+8; t+9; t+10; t+11} horizons, no other result can be rejected, meaning that the forecast would be right on average for all 86 others, as illustrated on *Figure 2*, displaying a pseudo forecast error ranging between -3% and 3%[11].

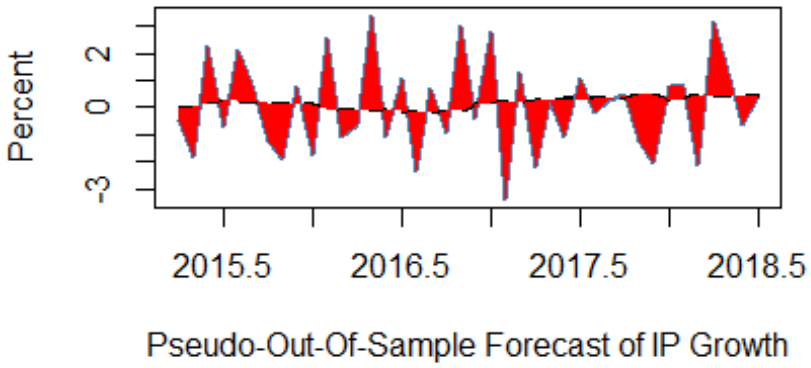

Figure 2: Pseudo forecast error of industrial production in South Korea full sample at the first horizon
(blue= Actual IP Growth Rate; black= Forecasted IP Growth Rate; Red= Pseudo Forecast Error)

A first quick analysis of the results would indicate a propension to precede the industrial production growth rate in the short-term, supporting the theory of Shareef and Shijin (2017) regarding the impact of fiscal deficit on the short-term interest rates. Moreover, it would seem that Asian countries tend to perform better than the others.

Nevertheless, as highlighted by Stock and Watson (1996), instability in the models lead to diminished performances, which could explain the previous results.

[10] The statistic value being lower than 0.001.
[11] Forecast including the intercept

Indeed, when looking at the results in *Table 4* and *Table* 5, obtained when accounting a break in the model thanks to a QLR test, different conclusions can be reached. Given the following break date, South Korea and Mexico will not be considered as the rule of thumb of 60 observations can no longer be respected:

- Brazil: 12/2008
- India: 09/2009
- Israel: 02/2006
- Mexico: 06/2014 (39 observations left in-sample)
- Russia: 01/2009
- South Korea: 10/2015 (33 observations left in-sample)

In-sample forecasts for those reduced sets indicate that two results only are left significant at 5%, Brazil's 9^{th} and 10^{th} horizons ahead, confirming partially at least Wheelock and Wohar's observation in 2009 about significant predictive power results when forecasting 6 to 12-months ahead. Enlarged to 10%, Brazil's 8^{th} horizon is the sole addition.

As for pseudo out-of-sample forecasts, all models built appear to be unable to reject the null hypothesis, that is are right on average, supporting the two authors' conclusion that the spread does hold predictive power.

Overall, if breaks do appear to take place during the Great Recession, no inference can be deducted regarding differences due to the political system, monetary policies nor the geolocation since results are relatively standardized across the set.

5.2. Inflation growth rate

Looking at results when forecasting the inflation growth rate presented in *Tables 7* to *12*, if they appear to be, on average, more significant than for the industrial production, here again, no consensus can be reached.

As a point of comparison to check for model improvements thanks to break date, outcomes were also computed for the full set. Much like before, full set results have to be treated carefully as they can be misleading. Nevertheless, more than anything, they would tend to support the idea that no common trend can be found across countries. Indeed, out of 198 results, 19 are significant at 5% or lower. Out of those, four happen to be at the 17^{th} horizon or more

(Brazil {t+18}, Russia {t+17}, South Korea {t+17; t+18}), three before the 6th month-ahead (Kenya {t+4; t+5} and Uruguay {t+4}), and 12, as suspected by Wheelock and Wohar (2009), within the 6 to 12-months ahead bracket. Kenya's outcomes appear to bear predictive power when forecasting 6, 7 and 9 months ahead, as Russian 10 to 12- months ahead do, besides Uruguay 12th period and South Korea 8th and to 12th horizon. Enlarged to 10% as a threshold for significativity, 20 more forecasts can be declared as significant, for the most part[12] taking place after the 12th month-ahead. Apart from Mexico for which the 10th period is now considered significant, the said enlargement includes additional horizons for the countries already found to bear some significativity depending on the forecast considered.

Regarding the (Newey-West) standard error, much like for the industrial production, results, ranging from 0.001 to 0.265, appear to be stable both within one's outcomes, and when cross-checked with others.

Unlike precedingly, the Single Sample t-test in this case do reject most of the forecasts obtained (113). If, for China, the null hypothesis[13] can only be rejected at 5% or lower for observations at {t+2; t+3; t+4}, all 18 periods are rejected for India, Indonesia, Israel, Russia, and a majority of them for Kenya (15) and South Korea (8).

As for the stability in terms of mean forecasted error, it can only be observed within a nation's results as large variations can be observed from one country to another (going from -0.777 to 0.811).

It would appear that some geographical trends are, in the case of inflation, of influence, as illustrated by Asian countries only just above. Moreover, a strict control of its growth should, as explained, worsen the result. Nevertheless, out of the four countries, determined in this paper, applying a strict control of the inflation, results are mitigated as well, with Uruguay and Russia bearing some predictive power, and having some forecasts right on average, whereas Israel and South Africa are indeed, wrong overall.

If those inferences can be imagined, it is important to verify them as those discussed results occurred when working with the full set. As suspected most breaks do occur during the Great Recession. Indeed, out of 11 nations, 4 happen in 2008 and 2 in 2009. For the rest, Russia (2005), Mexico (2006), South Korea (2015) and Uruguay (2012) can be seen as outliers. The

[12] 9 out of 20

[13] Being the true mean equals zero

two latter, given the rule of thumb of 60 observations, are not analyzed with a break, having respectively 34 and 49 observations left only.

Furthermore, those dates, while it would seem to support the idea of a break taking place due to the economical context, in fact account mostly for model instabilities resulting from higher volatility, instead of a new monetary policy stand potentially leading to more unstable macroeconomic factors, improving forecast performances.

Undeniably, when looking for an unknown break date in South Africa, for which data start in 1981, the outcome, December 1986, does support what the literature has highlighted, being that of a more stable political context ever since the mid-80s, impacting the spread results. As a reminder, the test used, being a QLR, choosing the date presenting the maximum F-statistics, it appears safe to assume than despite the Great Recession, the mid-80s constitute a more important break. Furthermore, governments most likely closely kept an eye on inflation growth rate during the said recession, worsening the performances,

Again, as suspected, performances differ greatly from what was found with the full set. Truly, countries for which the term spread do bear predictive power are now Brazil, Israel and Kenya, already infirming the conclusions that could have been made based on the full set.

In-sample Granger results tend to indicate that, if the idea of accurate forecast between the 6^{th} and 12^{th}-period ahead still seems correct (Brazil {t+7}, Kenya {t+6; t+7; t+8}, Israel {t+11; t+12}), some longer period forecasts appear to be significant as well. Aside from Kenya 3^{rd},4^{th} and 5^{th}-month ahead, remaining significant periods are all included between the 13^{th} and 18^{th} horizon (Kenya 16^{th}, Israel 13^{th} to 17^{th}, Kenya 13^{th} to 18^{th}). 5 observations could be included as well if looking at a 10% threshold, Brazil's 8^{th}, 9^{th} and 10^{th} period as well as Israel 6^{th} and 18^{th}.

Besides, standard errors present a higher stability, now ranging from 0.011 to 0.155 across the whole set.

Surprisingly, pseudo out-of-sample performances still reject a non-negligible number of outcomes, with, for instance, all 18 periods of Russia and South Africa rejecting the t-test null hypothesis, hence indicating a forecast wrong on average. China as well as India and Indonesia also reject most of their results. Like previously, if standard errors in sample do show stability both for one's results and between all, pseudo out-of-sample, even if demonstrate a relative stability when analysing one's outcomes, vary greatly from one country to another.

Overall, in the case of inflation, results do indicate some geographical trends, regardless of the set considered, as illustrated by the reduced set t-test, rejected mostly in Asian countries.

Moreover, despite Israel knowing significant result but being at a more advanced stage economically speaking, it appears that the monetary regime, and more specifically the strict surveillance of inflation does influence the results. Indeed, South Africa and Russia are the only two countries rejecting the forecast righteousness at all horizons.

Finally, it appears that Great Recession or not, the structural break of the mid-80s analysed and described in the literature holds more importance in terms of structural change than any date picked during the 2000s, probably reflecting more model instabilities than new monetary policies.

5.3. Unemployment growth rate

Analysing the forecast results of the spread concerning the unemployment growth rate bring little new evidence supporting the thesis and main hypothesis behind this paper. For this macroeconomic factor, data were retrieved for 5 countries only: Brazil, Mexico, Russia, South Korea and Uruguay.

If full set results will not be analysed here for the reasons explained previously, it is important to note that Mexico, given the smaller number of data retrieved for the unemployment rate, does not respect the rule of thumb of 60 observations when accounting for a structural break (48). Even if having to be treated carefully, it is interesting to know that none of its performances appeared to be significant in-sample despite a forecast right on average pseudo out-of-sample, alike the other nations as developed further.

The breaks obtained thanks to the QLR test are the following:

- Brazil: 12/2008
- Mexico: 12/2005
- Russia: 01/2009
- South Korea: 10/2006
- Uruguay: 12/2006

Once again, it would appear that even if the South African case demonstrated that despite structural changes occurring around or during the Great Recession, the mid-1980s one held more significativity, breaks can indeed be found in the suspected period.

As for the performances, as suspected earlier in the paper, it seems that the political system/the structure of the economy do impact them when forecasting unemployment rate. Indeed, if the inner workings of Russia will not be developed here, out of the 4 nations remaining, it is the only one knowing some significant horizons when considering Granger outcomes, {t+1; t+16; t+17}, as well as {t+2} if accepting in-between 5 and 10% values.

Furthermore, it is again the sole country rejecting the null hypothesis for the Single Sample t-test, realized pseudo out-of-sample, at {t+14}.

Interestingly enough, we can see that the stability demonstrated in terms of standard errors for one's outcomes as well as crossed did worsen, as well as the intern stability for the mean forecasted errors that entirely disappeared, showing volatility both within and across sets[14].

If no conclusion can be drawn regarding the influence of strict monetary regime, it appears that the Great Recession did result in structural change but not important enough to improve the results, whereas differences in political regime seem to be the main driver of variation.

[14] Read here the full database or each country set

6. Conclusion

Overall, if results obtained do not shed a new light to the *"stylized fact"* that is the term spread, reinforcing the idea of "no-consensus" even (variation from country to country,…), it still seems safe to assume that it does bear usefulness in forecasting macroeconomic factors, as most of the Single Sample t-test outcomes illustrate, being unable to reject their respective null hypothesis, hence acknowledging the model considered as right on average, regardless of the horizon and the region considered.

However, in terms of the empirical slope predictive power, not enough evidence can be gathered as to support the thesis that the Great Recession did result in a greater instability in terms of macroeconomic indicators, leading to better forecast performances as suggested by Wheelock and Wohar (2009).

Indeed, when looking at reduced sets results, regardless of the growth rate, a mere 7% only is found to "G-cause", in-sample at least, the explained the variable[15]. Looking at the inflation growth rate only, this percentage peaks around 10%.

Furthermore, this paper found evidence of geographical trends, as illustrated by Asian countries knowing significant results when put in relationship with inflation growth, political and economical differences are also major factor of influence, as demonstrated by Russia's results when considering the unemployment growth rate.

This does support the *"intertemporal consumption theories"*, reflecting the influence of external factors to the monetary policies.

As analyzed through Israel, Russia, South Africa and Uruguay, the strict inflation control assumption, allegedly worsening forecast performances, does seem to hold despite Israel significant results forecasting inflation. Indeed, if Israel's economy is most likely developed enough not to suffer too much of a loss in terms of predictive power, South Africa and Russia's outcomes tend to indicate that it does in fact have an impact.

Nonetheless, no further evidence could be found, hence requiring taking this inference cautiously.

[15] Read one of the three growth rates

Finally, main hypothesis directly drippling down from the thesis, the research of a structural break due to the Great Recession, while conclusive, do not bring new conclusion to the literature as, as demonstrated through the South African case, changes occurring in mid-1980s, reflecting the higher stability of macroeconomic factors and impacting the term spread predictive power, a QLR test indicates that the said South African break (1986) does present a higher F-statistics than whichever could take place during the period studied. Therefore, there are not enough evidence to conclude that the considered recession impacted enough the said factors, hence improving the predictive power of the spread, and that a new structural break should be considered by the researchers going forward.

As this report wanted to analyse the empirical slope predictive power in its simplest form, the model used could and should further be improved by future researchers in order to check the presented results through something closer to the reality, by fully decomposing the term spread since it has been proven that control variables do not add significativity to it.

Furthermore, not having been able to retrieve common data for all nations, it would be interesting to add them all but more specifically: China's, Israel's, Uruguay's, Russia's and South Africa's in order to further confirm or infirm the assumptions underlying this paper.

At last, not being an expert on the topic, mistakes were probably made, and room for improvements exists. Having found those two quotes of particular interest, I strongly suggest to whoever future researcher to keep them in mind as to settle on a choice instead of diversifying into multiple distinctive paths of modelization, etc.

"All models are wrong. Some are useful."

George Box.

"The only function of economic forecasting is to make astrology look respectable."

John Kenneth Galbraith.

7. Bibliography

- Arnold, M., Hanck, C., Gerber, A., & Schmelzer, M. (2019). *Introduction to Econometrics with R.* Syllabus, Department of Business Administration and Economics of the University of Duisburg-Essen.
- Banco Central del Uruguay. (2019). *A stability-oriented monetary policy.* Retrieved on https://www.bcu.gub.uy/Ingles/Paginas/Monetary-Policy.aspx
- Bank of Israel. (2019). *Conducting Monetary Policy.* Retrieved on https://www.boi.org.il/en/MonetaryPolicy/MonetaryPolicyFramework/Pages/Default.aspx
- Bauer, M.D., & Mertens, T.M. (2018). Economic Forecasts with the Yield Curve. *Federal Reserve Bank of San Francisco Economic Research,* Economic Letter No. 2018-07, 1-5. Retrieved on https://www.frbsf.org/economic-research/publications/economic-letter/2018/march/economic-forecasts-with-yield-curve/
- Benati, L., & Goodhart, C. (2008). Investigating Time-Variation in the Marginal Predictive Power of the Yield Spread. *Journal of Economic Dynamics and Control,* 32(4), 1236-1276.
- Berk, J.M., & van Bergeijk, P.A.G. (2001). On the Information Content of the Yield Curve: Lessons for the Eurosystem?. *Kredit und Kapital,* 1, 28-47.
- Bordo, M.D., & Haubrich, J.G. (2004). The Yield Curve, Recessions, and the Credibility of the Monetary Regime: Long-Run Evidence, 1875-1997. *National Bureau of Economic Research,* NBER Working Paper No. 10431, 1-36. Retrieved on https://www.nber.org/papers/w10431.pdf?new_window=1.
- Bordo, M.D., & Haubrich, J.G. (2008). The Yield Curve as a Predictor of Growth: Long-Run Evidence, 1875-1997. *Review of Economics and Statistics,* 90(1), 182-185.
- Chen, J. (2019). *Term Structure of Interest Rates.* Retrieved on Investopedia https://www.investopedia.com/terms/t/termstructure.asp
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. J. (1990). STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics, 6*(1), 3–33.

- D'Agostino, A., Giannone, D., & Surico, P. (2006). (Un)predictability and Macroeconomic Stability. *European Central Bank,* ECB Working Paper No. 605, 1-43. Retrieved on https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp605.pdf
- Davis, E.P., & Fagan, G. (1997). Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in EU Countries?. *Journal of Applied Econometrics,* 12(6), 701-714.
- Dewachter, H., Iania, L., & Lyrio, M. (2014). Information in the Yield Curve: A Macro-Finance approach. *National Bank of Belgium,* Working Paper No. 254, 1-29. Retrieved on https://www.nbb.be/doc/ts/publications/wp/wp254en.pdf
- Dotsey, M. (1998). The Predictive Content of the Interest Rate Yield Spread for Future Economic Growth. *Federal Reserve Bank of Richmond Economic Quarterly*, 84(3), 31-51. Retrieved on https://www.richmondfed.org/~/media/richmondfedorg/publications/research/economic_quarterly/1998/summer/pdf/dotsey.pdf
- Estrella, A., Rodrigues, A.P., & Schich, S. (2003). How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States. *Review of Economics and Statistics,* 85(3), 629-644.
- Estrella, A. (2005). Why Does the Yield Curve Predict Output and Inflation?. *Federal Reserve Bank of New York Economic Journal,* 115(issue 505), 722-744. Retrieved on https://doi.org/10.1111/j.1468-0297.2005.01017.x
- Estrella, A., & Trubin, M.R. (2006). The Yield Curve as a Leading Indicator: Some Practical Issues. *Federal Reserve Bank of New York Current Issues in Economics and Finance*, 12(5), 1-8. Retrieved on https://www.newyorkfed.org/medialibrary/media/research/current_issues/ci12-5.pdf
- Feroli, M. (2004). Monetary Policy and the Information Content of the Yield Spread. *Topics in Macroeconomics*, 4(1), Article 13.
- Hännikäinen, J. (2016). When Does the Yield Curve Contain Predictive Power? Evidence from a Data-Rich Environment. *Tampere Economic Working Paper*, Working Paper No. 103, 2-25. Retrieved on https://doi.org/10.1016/j.ijforecast.2017.05.006

- Harvey, C.R. (1988). The Real Term Structure and Consumption Growth. *Journal of Financial Economics*, 22(2), 305-333.
- Harvey, C.R. (1991). The Term Structure and World Economic Growth. *Journal of Fixed Income*, 1(1), 7-19.
- Haubrich, J.G., & Dombrosky, A.M. (1996). Predicting Real Growth Using the Yield Curve. *Federal Reserve Bank of Cleveland Economic Review,* 32(1), 26-35. Retrieved on https://www.clevelandfed.org/newsroom-and-events/publications/discontinued-publications/economic-review/1996-economic-review/er-1996q1-predicting-real-growth-using-the-yield-curve.aspx
- Hu, Z. (1993). The Yield Curve and Real Economic Activity. *IMF Staff Papers,* 40(4), 781-806.
- Kent State University. (2019). *One Sample t Test.* Retrieved on https://libguides.library.kent.edu/SPSS/OneSampletTest
- Philips, M., & Grocer, S. (2019, 31 May). The Bond Market Is Trying To Tell Us Something (Worry). *The New York Times*, 1. Retrieved on https://www.nytimes.com/2019/05/30/business/bond-yield-curve-recession.html
- Schuiling, I. (2018). *International Marketing Management.* Syllabus, Louvain School of Management of the Université Catholique de Louvain.
- Shareef, H.K., & Shijin, S. (2017). The term structure of interest rates and macroeconomic factors: evidence from Indian financial market. *Borsa Istanbul Review,* 17(4), 249-256.
- South African Reserve Bank. (2019). *Monetary Policy.* Retrieved on https://www.resbank.co.za/MonetaryPolicy/Pages/MonetaryPolicy-Home.aspx
- Stock, J.H., & Watson, M.W. (1996). Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business and Economic Statistics,* 14(1), 11-30.
- The Central Bank of the Russian Federation. (2019). *Monetary Policy.* Retrieved on https://www.cbr.ru/eng/DKP/
- TradingEconomics. (2019). *Country list government debt to GDP.* Retrieved on https://tradingeconomics.com/country-list/government-debt-to-gdp

- Wheelock, D.C., & Wohar, M.E. (2009). Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature. *Federal Reserve Bank of St. Louis Review,* 91(5, Part 1), 419-440. Retrieved on https://research.stlouisfed.org/publications/review/2009/09/01/can-the-term-spread-predict-output-growth-and-recessions-a-survey-of-the-literature