NOTAS ECONÓMICAS 57

ARTICLES

Spyridon Boikos / Eirini Makantasi / Theodore Panagiotidis
A Macroeconomic Uncertainty Index for European Countries

Marcelo Santos / Susana Garrido / Tiago Sequeira
Management and Human Capital Employment: An Overlooked Relationship

LETTERS

Tomé Lima / Helder Sebastião
Native Market Factors for Pricing Cryptocurrencies

Pedro Bação / José Diogo Silva
Inflation in Portugal Through the Lens of the Fair Model

Imprensa da Universidade de Coimbra
Coimbra University Press

Faculdade de Economia da
Universidade de Coimbra

C O N Ó M I C A S 57

N O T A S

Informações aos autores (Submission guidelines)
https://impactum-journals.uc.pt/notaseconomicas/about/submissions

Publicação resumida e indexada por (Indexing and Abstracting)
EconLit
RePEc
REDIB
latindex

Design
Imprensa da Universidade de Coimbra

Execução gráfica
Talento & Tradição

Tiragem 200 exemplares

Propriedade
Faculdade de Economia
Universidade de Coimbra
Av. Dias da Silva, 165
3004-512 Coimbra

Apoios (Sponsors)
Esta publicação é financiada por fundos nacionais através da
FCT – Fundação para a Ciência e a Tecnologia, I.P.,
no âmbito do projeto UIDB/05037/2020

FCT Fundação para a Ciência e a Tecnologia

Número avulso
Portugal 7,5 €
Outros países 15 €

Condições de assinatura bienal
Portugal: 20 €
Outros países: 40 €

ISSN 0872-4733

ISSN Digital 2183-203X

DOI https://doi.org/10.14195/2183-203X_57

Depósito Legal 64932/93

Contents

Articles

Letters

Spyridon Boikos
Department of Economics
University of Macedonia, Greece
boikos@uom.edu.gr
orcid.org/ 0000-0003-4907-8223
Eirini Makantasi
Department of Economics
University of Macedonia, Greece
mae23005@uom.edu.gr
Theodore Panagiotidis
Department of Economics
University of Macedonia, Greece.
tpanag@uom.edu.gr
orcid.org/ 0000-0002-5328-3054

DOI: https://doi.org/10.14195/2183-203X_57_1

Macroeconomic Uncertainty Indices for European Countries

Índices de Incerteza Macroeconómica para Países Europeus

Spyridon Boikos
Eirini Makantasi
Theodore Panagiotidis

Received for publication: October 1, 2023
Revision accepted for publication: December 12, 2023

ABSTRACT
The present article seeks to develop a macroeconomic uncertainty index for the EU Member States based on Google Trends for a period of fifteen years (from January 2008 to December 2022). Monthly data were collected for the 12 countries for four different word-terms, as well as for unemployment rate, inflation and the 10-year Government Bond yield. For simplifying the research the keywords searched were in English and were not translated into the countries' own languages. Our findings were then compared to existing uncertainty indices. Lastly, we employed Impulse Response Functions (IRFs) with the existing economic indicators to highlight the effect that one standard deviation shock on the uncertainty index has on all three indicators and its ability to accurately depict the future precariousness of the country.
Keywords: Uncertainty; Google trends; European uncertainty index.

JEL:Classification: C32; E32

1. Introduction

In this paper we try to create a macroeconomic uncertainty index for each of the 12 core Eurozone countries (the countries in which euro currency went initially into circulation on the 1st of January 2002). The creation of the macroeconomic uncertainty index is based on data gained from Google Trends. Then we are interested in checking the impact of the uncertainty index of each country through Impulse Response Functions (IRFs) on three specific macroeconomic variables: unemployment rate, inflation and the 10-year Government Bond yield. Finally, we try to check with the tool of (IRFs) the effect of the uncertainty index of Germany (the biggest economy in EU and in Eurozone) both on the individual uncertainty index of each country and on the three macroeconomic variables of interest. The Google Trends tracks the most popular Google Search terms across various geographies and languages. In our paper we have used four common words all in the English language in order to create the uncertainty index. The dataset starts at January 2008 and ends at December 2022, which implies 15 years which includes the period of debt crisis for some countries such as Greece, Ireland, the covid pandemic and the war in Ukraine. All these events have created bank crisis as in Greece, a push both in energy and home prices. All the previous elements may increase income inequality and the minimum wage is of paramount importance for the wellbeing of the society. Under the previous justification and by taking into account the proposals from the literature, we have decided to use the following four words in order to construct the uncertainty index for each country: bank crisis, energy price, home price and minimum wages.

The Uncertainty Index was constructed utilizing Google Trends, obtaining monthly data for all 12 core Eurozone countries based on four benchmark words. The 12 core Eurozone countries in alphabetical order are the following: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal and Spain. The validity of our uncertainty index was assessed against the established Economic Policy Uncertainty index (Baker et al., 2016) and the Consumer Confidence Index. Utilizing the STATA econometric program, Vector Autoregressive (VAR) models were conducted for each country, followed by the depiction of Impulse Response Function (IRF) graphs illustrating the impact of a one standard deviation shock in country uncertainty on economic indicators such as unemployment rate, inflation, and long-term government bond yield. The structure of the paper is the following: in section 2 we provide the literature review, in section 3 we explain the construction of the uncertainty index for each country by using data from google trends and we provide the empirical results for each country. In the last section as usual there are the conclusions.

2. Literature Review

In the literature review, the significance of textual analysis has been extensively documented. Examples include Dergiades et al. (2015), Milas et al. (2021), and Bampinas et al (2019). Schütze (2020) employs Google Trends subject searches to develop an uncertainty index applicable to countries where Google operates. The uncertainty indicator generated in this study consistently yields statistically significant results higher than those of the EPU on average. The study concludes that Google Trends serves as an effective instrument for obtaining timely information on economic participants' uncertainty. Notably, the primary enhancement lies in the independence of this uncertainty proxy from language.

Castelnuovo and Tran (2017) utilized publicly available, real-time Google Trends data to devise uncertainty indices for both the United States and Australia. The terms employed in crafting the uncertainty index were sourced from economic documents such as the Federal Reserve Beige Book for the US and the Reserve Bank Monetary Policy Statement for Australia. The authors demonstrate that several other proxies for uncertainty applicable to these two nations exhibit favorable correlations with the Google Trends Uncertainty (GTU) indices they developed, including VXO as used by Bloom (2009) and the EPU index constructed by Baker et al. (2016). Through investigations using VAR, it was revealed that GTU shocks in the United States exert a statistically and economically substantial impact on the dynamics of unemployment. Conversely, GTU shocks were found to have a significantly smaller and less significant impact on Australian unemployment dynamics compared to shocks related to monetary policy.

Donadelli (2015) proposed three distinct metrics of policy-related uncertainty by using the frequency of Google searches for terms such as "US stock market", "US politics", and "US Fed". He found out that a Google search-based uncertainty shock significantly and negatively affects US macroeconomic conditions in a VAR environment. Specifically, it leads to reductions in industrial production, consumer confidence, equity prices, long-term rates, and consumer credit. Another finding of this paper is that uncertainty shocks contribute to an increase in the unemployment rate. The empirical results suggest that a surge in the number of online searches related to themes linked to economic policy signals rising uncertainty. The proposed Google-search-based measures align well with common policy-related uncertainty indicators, such as the EPU index developed by Baker et al. (2016) and the VIX (Volatility Index).

Moore (2017) developed a monthly indicator of economic uncertainty for Australia. During the global financial crisis, economic uncertainty reached unprecedented levels and persisted until 2013. He finds out that the economic uncertainty index tends to rise faster than it falls, influenced by both domestic and international factors, and is particularly pronounced around recessions, elections, monetary policy shocks, and significant geopolitical events. He concludes that it hampers investment and job creation, consistent with the real options' channel of uncertainty. Similarly, akin to the 'precautionary savings' channel of uncertainty, uncertainty raises the household saving ratio and reduces consumption growth for durable goods.

Albert and Fernández (2018) utilize data spanning from January 2001 to June 2018 to employ a SVAR technique with sign restrictions. The aim is to estimate the effects of economic uncertainty shocks on key macroeconomic variables in Spain. The authors investigate both short-term and long-lasting shocks associated with economic uncertainty. Furthermore, they isolate uncertainty shocks originating solely from political sources to discern potential variations in their impact. Their findings suggest that increases in economic and political uncertainty lead to higher unemployment rates and decreases in both company and consumer confidence, the IBEX 35 Index, and industrial production. Moreover, these adverse effects of uncertainty persist over a prolonged period, especially in the cases of industrial output and unemployment. Based on these results, the authors conclude that economic uncertainty shocks exert a significant negative impact on the Spanish economy. Moreover, the research suggests that political stability is crucial in mitigating uncertainty and achieving improved economic outcomes.

Bontempi et al. (2016) paper tries to investigate the impact of uncertainty index, which is constructed by internet searches, on the economic cycle. Moreover, they compare the macroeconomic consequences of various uncertainty indices. The findings suggest that uncertainty shocks, at times, convey relevant information regarding people's perceptions of uncertainty sooner than other indices. Bilgin et al. (2019) measures the level of economic and financial uncertainty in Turkey. The uncertainty index is measured with the use of internet search-based method and it provides the 'Turkish Economic and Financial Uncertainty Index' (TEFUI). They have used real-time monthly Google Trends data for the period from January 2004 to December 2018. In order to create the baseline TEFUI, the paper takes into account more than 400 possible terms. The results of the Vector Autoregression models, Impulse-Response shocks and correlation analysis showed that the TEFUI is substantially correlated with a number of domestic economic uncertainty indicators and global uncertainty indices.

Kropiński and Anholcer (2022) explore the correlations between the WIG20 index and phrases associated with economic policy uncertainty (EPU) measured through Google Trends search index. The examination covers two distinct timeframes: January 2015 to December 2019 and June 2016 to May 2021, allowing differentiation between a period of relative stability and the economic shock induced by the COVID-19 epidemic crisis and subsequent government-imposed restrictions. For their empirical analysis it is used a bivariate VAR model. The study found that twelve EPU-related keywords exhibited a stronger empirical association with changes in the WIG20 index during the post-COVID era compared to six terms in the pre-COVID period. Moreover, the severity of reversal relations increased notably throughout the post-COVID period.

Zayed et al. (2023) conducted a scoping review aiming to provide an overview of Google Trends' role as a monitoring and forecasting tool for the COVID-19 pandemic. The study focused on original English-language peer-reviewed research publications on the COVID-19 pandemic from 2020 that utilized Google Trends as a search engine. Articles not detailing the use of Google Trends during the COVID-19 epidemic, written in languages other than English, or available solely in abstract form were excluded. A total of 81 papers meeting the inclusion criteria were included, covering the first year following the emergence of the crisis. The findings suggested that health authorities could benefit from utilizing Google Trends to plan and manage pandemics earlier.

Bulczak (2021) with the utilization of Google Trends data tries to improve the real estate market forecasting. Online searches provide valuable information that precedes financial decisions. This study delves into the potential of Google search engine data in forecasting real estate markets. The findings indicate that Google data could serve as an additional source of insight for investors and decision-makers. Google Trends data has been identified as a reliable indicator of real estate market pricing and sales volume. However, Limnios and You (2021) investigate the use of Google Trends data to complement linear pricing models for the housing market, commonly employed in literature. They found that augmenting models with Google Trends data did not significantly enhance their predictive abilities.

Ettredge et al. (2005) highlight the promising potential of web-based search data for forecasting macroeconomic statistics. Through the analysis of the vast amount of data generated by internet search activity, researchers gained valuable insights into consumer attitudes and behavior. Hayford (2000) demonstrates that concern about future unemployment, which serves as a proxy for uncertainty regarding future actual economic activity, rises with inflation as well as inflation uncertainty itself. His results show that a temporary slowdown in production growth occurs when both inflation uncertainty and unemployment uncertainty rise. Further impulse response functions illustrate that the impacts of inflation and unemployment uncertainty on real GDP growth are of similar magnitude.

3. **Data and Empirical Results**

By using Google Trends at monthly basis from January 2008 to December 2022 we have created the Uncertainty Index (UI) for each core country of the Eurozone. To keep the research simple, the terms that were selected and examined were in English rather than being translated into the native tongue of each nation. In order to create the uncertainty index, we selected terms and phrases that, during times of increasing uncertainty, people would be most likely to use to search for information on Google, the most widely used search engine worldwide. "Minimum wage," "energy price," "bank crisis," and "home price" were these four terms. We think that these four terms can capture better the uncertainty of the period since in this period the following types of crises have appeared: debt crisis in Greece, Covid pandemic crisis and the Ukrainian war. All these words can capture mainly uncertainty which is more related with increasing inflation, which is something that at the moment both Eurozone and in general the whole world is facing.

Instead of looking at the total number of searches, Google Trends data shows us the percentage of searches on a particular topic relative to all searches made at that time and place. Since Google Trends data is derived from an impartial, random sample of Google searches, we gathered the data for the words under investigation for each of the 12 countries on the same day, even though the results change daily. For the sake of simplicity in our analysis, we rounded all values that came close to 1 for each of the four terms we looked into. Greece was our primary focus, so we used Google Trends data to establish it as the benchmark country. We collected data for each country using the same four terms, and we added the term "home price" for Greece as the fifth search. By using this technique, we were able to rescale the required number of countries, allowing the index to accept values up to 100 and to be comparable the uncertainty index between each country. The detailed exploration of the methodology for the construction of the uncertainty index is presented in Castelnuovo and Tran (2017).

We gathered monthly data for the "unemployment rate," "inflation," and "long-term government bond yield 10 year" for each country from January 2008 to December 2022 after compiling the data and creating our monthly Uncertainty Index (UI).

Next, we looked at the relationship between these economic indicators and our Uncertainty Index. The Federal Reserve Economic Data (FRED) St. Louis FED website served as our primary source of data for these variables. We first determined the correlation between our Uncertainty Index and the economic indicators for each country. Next, we determined the correlation between our Uncertainty Index and the Consumer Confidence Index (CCI) and the widely used Economic Policy Uncertainty Index (EPU), as reported by Baker et al. (2016). These indices are widely used in numerous research fields and are easily and freely accessible via their websites. Only the following countries' EPU data could be located: Belgium, France, Germany, Greece, Ireland, Italy, Netherlands, and Spain. For each correlation, a significance test was conducted to ensure that the values obtained were legitimate.

In addition, we used Vector Autoregressive (VAR) models to generate Impulse Response Function (IRF) graphs. This enabled us to assess the relationship between our Uncertainty Index and the three variables we used (inflation, unemployment rate, and 10-year government bond yield). First differences were taken whenever necessary, and tests for unit-roots, such as Phillips Perron and Augmented Dickey Fuller, were utilized. The ideal lags for the VAR model have been determined. We began with 12 lags for each country, which translates to 12 months when we use monthly data, and we eventually reached 2 lags. In order to determine the ideal lag, we needed the majority of the tests to display optimal lags, all eigenvalues to fall inside the unit circle so that the VAR could meet the stability requirement, and the second lag to be larger than 0.05 in order to remove autocorrelation. The unit root tests, optimal lag tests, and VAR results are not provided here, but they are available upon request.

We were able to examine how one standard deviation shock affected the nation's level of uncertainty regarding its economic indicators by using Impulse Response Function (IRF) graphs. Twelve periods of forecasting were set for each IRF. Additionally, we investigated the potential impact of Germany's robust economy's level of uncertainty on the economic indicators and uncertainty of other European nations. The following section provides a brief presentation of the findings. Each country is shown and discussed independently, with numerous graphs and tables included.

3.1. Austria

The first graph for each country shows the variation of the uncertainty index together with the variation over time for the endogenous variable which has the highest correlation with the uncertainty index. For all the Eurozone countries the endogenous variable which has the highest correlation with the variable of a country's uncertainty index is inflation. We depict with bold colored line the uncertainty index and without bold the variable of inflation. The first table presents the correlations between all the variables (endogenous and the variable of uncertainty). As we can observe in Table 1.1., the negative correlation between the Uncertainty Index and the unemployment rate does not line up with the theory, as high uncertainty may induce a drop in the number of vacancies and in the job finding rate, ultimately resulting in a rise in unemployment, but it is considered to be a very weak correlation.

Figure 1.1. Austria – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

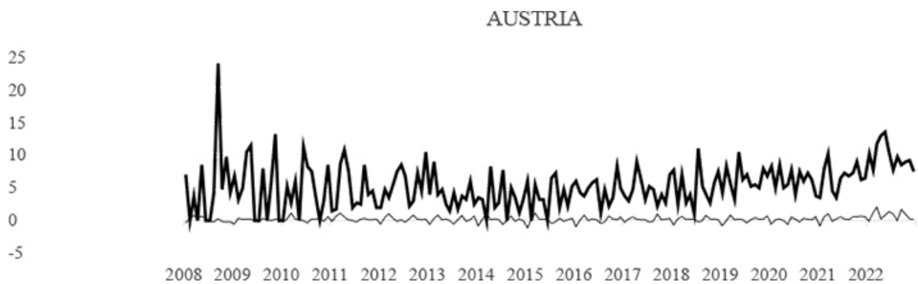


Table 1.1. Austria – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long- term government bond 10-year yield of the country

	Unemployment rate	LTGBY 10Y	Inflation	UI Austria
Unemployment rate	1			
LTGBY 10Y	-0.434895999	1		
Inflation	-0.16963067	-0.007424974	1	
UI Austria	-0.187562541	-0.061869347	0.137563423	1

We found data only for the Consumer Confidence Index (CCI) for Austria, which is relevant for examining the correlation between our Uncertainty Index and other existing uncertainty indices. The correlation coefficient of -0.2019 in Table 1.2.1 suggests a negative relationship between the two indices, which is in line with theory and is a desirable outcome. It indicates that a consumer's confidence decreases with each increase in uncertainty. The correlation's t-statistic is shown in Table 1.2.2 and was determined to be statistically significant at the 1% confidence level.

Table 1.2.1. Austria – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI AUT) and the Consumer Confidence Index (CCI AUT)

	UI AUT	CCI
UI AUT	1	
CCI AUT	-0.2019734	1

Table 1.2.2. Austria – t-statistic and p-value prices from correlations

	t-statistic	p-value
CCI/UI	-2.751364727	0.006548769

One of the largest economies in Europe, Germany, is compared in the following graph with Austria's Uncertainty Index, which is based on Google Trends. Throughout the analysis, the line that is not bold indicates each country's uncertainty index; in this case, it represents Austria's uncertainty index. The bold line represents Germany's uncertainty index. With a p-value of less than 0.01 and a t-statistic absolute value of 4.40, the table indicates a weak correlation of 0.3132, making it statistically significant at the 1% confidence level.

Figure 1.2. Austria – Google Trends based Uncertainty Index of Austria and Germany

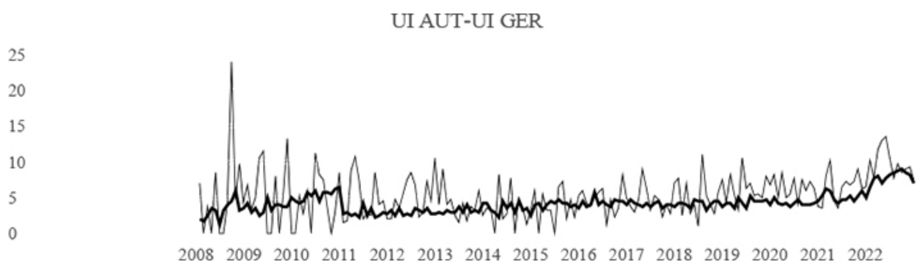


Table 1.3.1. Austria – Correlation between Austria's and Germany's Uncertainty Indices

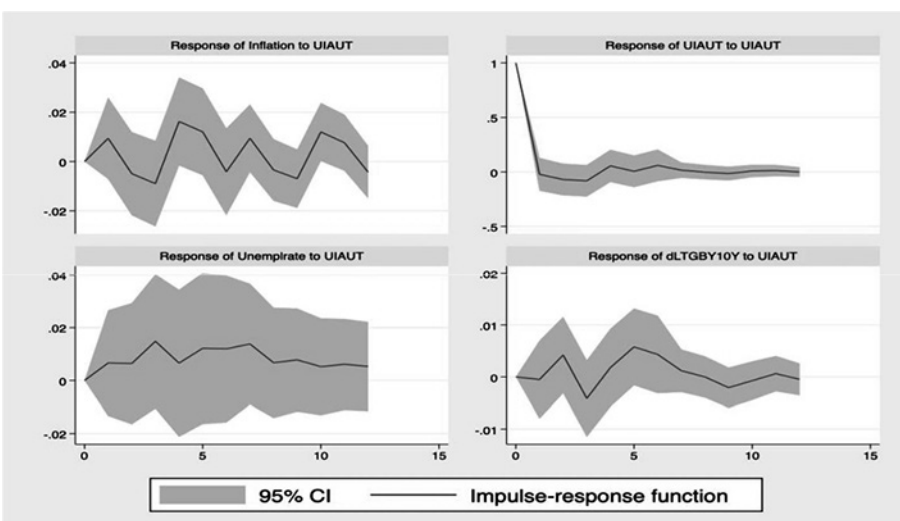
	UI GER	UI AUT
UI GER	1	
UI AUT	0.31325174	1

Table 1.3.2. Austria – t-statistic and p-value prices from correlations

	t-statistic	p-value
UI AUT/UI GER	4.400790665	1.85474E-05

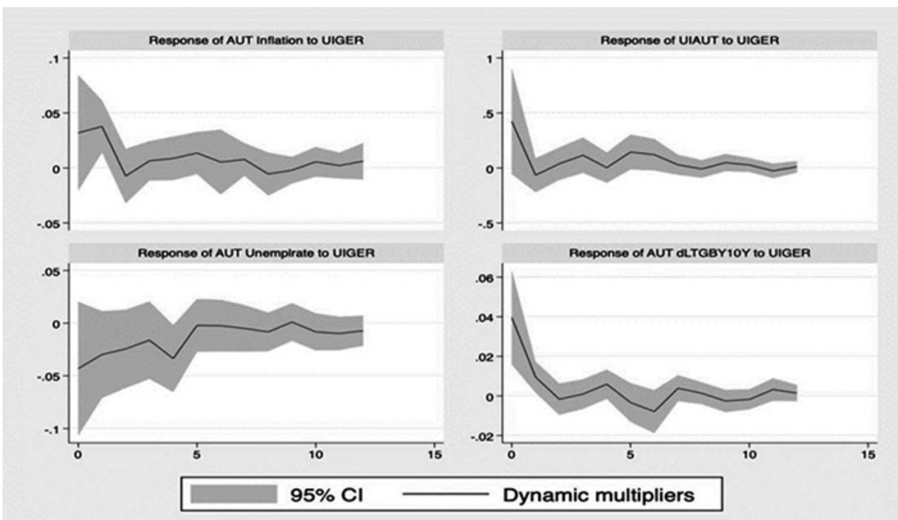
We begin with the conduction of the VAR model with reference to the Impulse Response Function (IRF) graphs. We ran the Phillips Perron and the Augmented Dickey-Fuller tests, with the alternative being that the variable was produced by a stationary process and the null hypothesis being that the variable contained a unit root. We have used solely stationary variables in all of the ensuing analyses. Since the variable LTGBY10Y for Austria was discovered to contain unit-root, the first differences were calculated. Six lags were determined to be the ideal values for the VAR model in order to satisfy every test. On the basis of this, IRF graphs were created. The first IRF graph shows the one-standard deviation impulse of our Austrian Uncertainty Index (UIAUT) to the country's dLTGBY10Y, the unemployment rate, the inflation rate, and the uncertainty index itself. The Austrian economic indices respond to such a shock in a minor and nearly insignificant way, with the inflationary response fluctuating between 0.01% and 0.04% in price level over the course of 12 periods. With respect to the UIAUT on UIAUT, the first shock occurs during the first period, but it soon fades away as the impact returns to 0.90% and then gradually drops to 0.85% after a year. In terms of the unemployment rate, the shock stays positive at the 0.01% level, and in terms of dLTGBY10Y, there is a slight decline during periods three and four, but the price level rises right away the following period.

Figure 1.3.1. Austria – Impulse Response Functions to a UIAUT shock. Sample: 2008M1 – 2022M12. VAR (6) estimated with an exogenous variable (UIGER). 95% confidence interval



A dynamic-multiplier function, which gauges the long-term effects of a unit increase in an exogenous variable on the endogenous variables, was employed for the second IRF graph. The Uncertainty Index for Germany is the exogenous variable in this scenario. The UIGER one-standard deviation impulse to the AUT Inflation, UIAUT, AUT Unemplrate, and AUT dLTGBY10Y is shown in the following graph. More of an indication response appears to be produced by the UIGER shock than by the UIAUT shock. The AUT Inflation spikes up to 0.03% very quickly before reverting to zero by the second period. Starting at 0.4%, the shock on UIAUT gradually decreases to approximately 0.1% and 0.02%. AUT Unemplrate shows a negative impact for all 12 periods, indicating that the shock to Germany's Uncertainty Index decreased Austria's unemployment rate. Finally, the AUT dLTGBY10Y fluctuates around zero after briefly remaining positive.

Figure 1.3.2. Austria – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (6) estimated with an exogenous variable (UIGER). 95% confidence interval



3.2. Belgium

Table 2.1 displays the correlations between the indicators. In this case, the unemployment rate and the Uncertainty Index have a moderately negative correlation (0f -0.50). Additionally, the weak 0.39 correlation between the 10-year government bond yield and the unemployment rate and the -0.22 correlation between inflation and unemployment rate are consistent with economic theory, since rising unemployment tends to drive down inflation and raise high-yield bond spreads.

Figure 2.1. Belgium – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

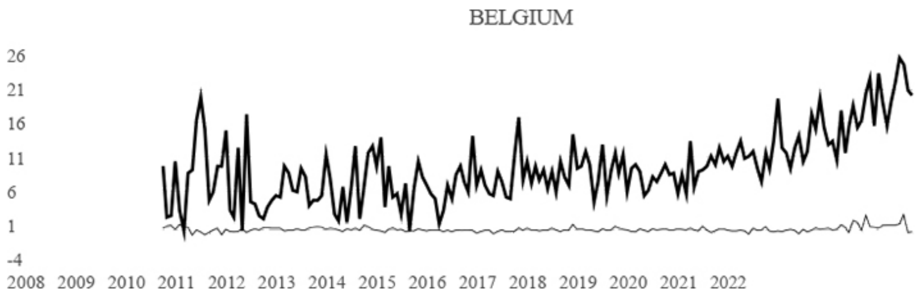


Table 2.1. Belgium – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield of the country

	Unemployment rate	Inflation	LTGBY 10Y	UI Belgium
Unemployment rate	1			
Inflation	-0.228293421	1		
LTGBY 10Y	0.390578592	0.016806041	1	
UI Belgium	-0.506568979	0.28430013	-0.340880245	1

We use Table 2.2.1. to examine the relationship between our Uncertainty Index and other available measures of uncertainty. According to theory, the correlation signals between UI BEL, CCI BEL, and EPU BEL are timely. One can always anticipate a negative correlation between the Consumer Confidence Index and an uncertainty index. The correlations' *t*-statistics and *p*-values are presented in Table 2.2.2, where they are statistically significant at the 1% confidence level.

Table 2.2.1. Belgium – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI BEL), the Consumer Confidence Index (CCI BEL) and the Economic Policy Uncertainty index (EPU BEL)

	UI BEL	CCI BEL	EPU BEL
UI BEL	1		
CCI BEL	-0.3171053	1	
EPU BEL	0.36288483	-0.4453187	1

Table 2.2.2 Belgium – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
CCI/UI	-4.460940427	1.44281E-05
EPU/UI	5.195654514	5.54546E-07
CCI/EPU	-6.635556658	3.75682E-10

It is noteworthy that the Uncertainty Index for each nation exhibits distinct patterns with respect to the duration of the COVID-19 pandemic crisis. From 2019 to 2022, the UIBEL nearly doubles, while the UIGER marginally rises. The correlation between UIBEL and UIGER is presented in Table 2.3.1. a *t*-statistic of 10.62, a *p*-value well below 0.01, and a positive and robust correlation of 0.622 between the nations.

Figure 2.2. Belgium – Google Trends based Uncertainty Index of Belgium and Germany

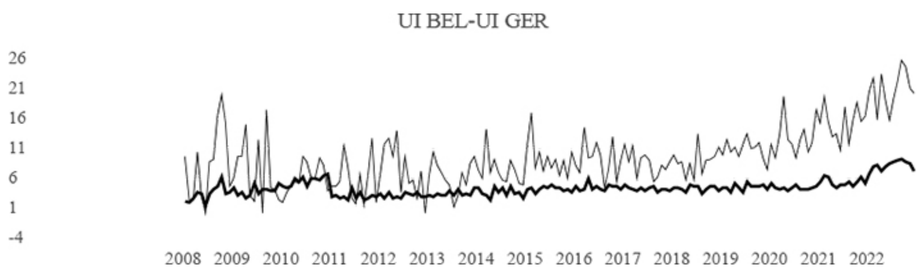


Table 2.3.1. Belgium – Correlation between Belgium's and Germany's Uncertainty Indices

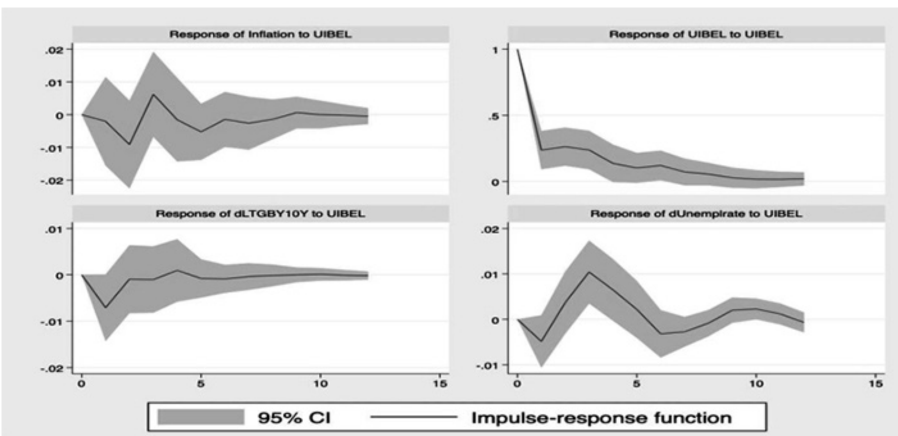
	UI GER	UI BEL
UI GER	1	
UI BEL	0.62294019	1

Table 2.3.2. Belgium – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
UI BEL/UI GER	10.62429146	9.89426E-21

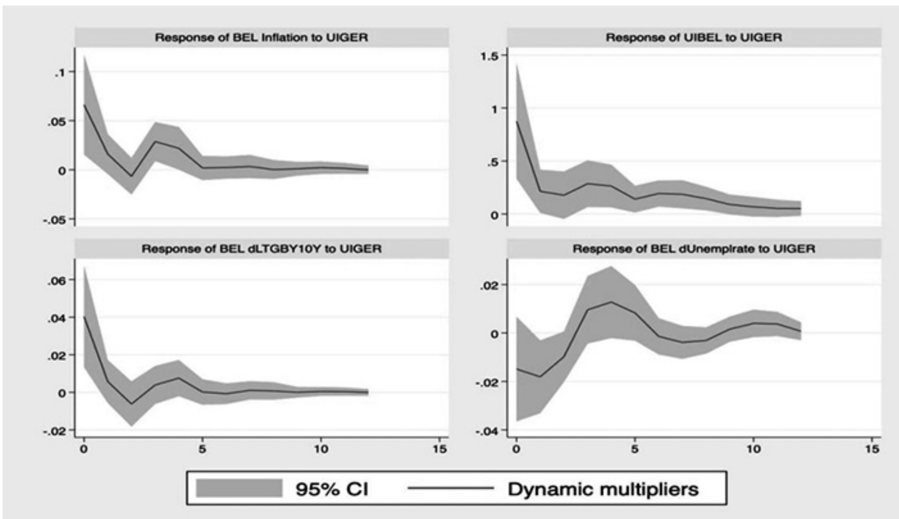
Four lags were required for the VAR model. The first IRF graph shows the one-standard deviation impact of our Uncertainty Index of Belgium (UIBEL) on the nation's dLTGBY10Y, dUnemplrate, inflation, and uncertainty index itself. Despite three brief periods of slight improvement, it appears that the inflation's price levels stayed mostly negative. In reference to the UIBEL, it peaked at 0.26% during the second period and then varied between 10% and zero until the sixth period, when it eventually died out and converged to zero. The only variable that continued to be negative over the course of the 12 periods was the dLTGBY10Y. Currently, dUnemplrate was negative only during the first period before rising to a peak price of 0.01%.

Figure 2.3.1. Belgium – Impulse Response Functions to a UIBEL shock. Sample: 2008M1 – 2022M12. VAR (4) estimated with an exogenous variable (UIGER). 95% confidence interval



A dynamic-multiplier function was utilized to assess the long-term effects of a unit increase in an exogenous variable on the endogenous variables in the second IRF graph. Germany's Uncertainty Index is the exogenous variable, and the BEL Inflation, UIBEL, BEL dUnemplrate, and BELdLTGBY10Y are the endogenous variables. Once more, it seems that the UIGER has a bigger impact on Belgian economic metrics than the UIBEL. The graphs for BEL dUnemplrate, UIBEL, and BEL Inflation appear to be similar. Only UIBEL manages to stay positive over the course of all 12 periods, with all three starting out positively in the first period and the first few months. Around the third period, BEL Inflation and BEL dUnemplrate both turn negative and briefly turn positive before approaching zero. When comparing BEL dLTGBY10Y to UIBEL, which peaked at 0.007%, it is evident that UIGER has a larger influence because its range spans almost -0.02% to 0.01%.

Figure 2.3.2. Belgium – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (4) estimated with an exogenous variable (UIGER). 95% confidence interval



3.3. Finland

The inflation rate, unemployment rate, 10-year government bond yield, and our calculated Uncertainty Index for Finland are all shown in the graph below. The 2008 financial crisis is when the Uncertainty Index reaches its highest value. The COVID-19 crisis aftermath of 2021–2022 also represents a period of increased uncertainty. For a considerable amount of time, the unemployment rate seems to be constant, reaching its 2008 level in little more than a decade. Table 3.1 reports all correlations as weak and negative apart from inflation.

Figure 3.1. Finland – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

Table 3.1. Finland – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Finland
Unemployment rate	1			
LTGBY 10Y	-0.133259754	1		
Inflation	-0.260016808	0.09665814	1	
UI Finland	-0.233854465	-0.215197104	0.19575778	1

Once more, the Consumer Confidence Index for Finland was discovered, but the Economic Policy Uncertainty Index had no data. In this scenario, the desirable result is a negative correlation. At the 5% confidence level, the correlation of -0.190 is statistically significant.

Table 3.2.1. Finland – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI FIN) and the Consumer Confidence Index (CCI FIN)

	UI FIN	CCI FIN
UI FIN	1	
CCI FIN	-0.1905386	1

Table 3.2.2. Finland – t-statistic and p-value prices from correlations

	t-statistic	p-value
CCI/UI	-2.589542738	0.010405644

The comparison of the two Uncertainty Indices between the two nations is shown in Figure 3.2. The p-value of the t-statistic is equal to 1, indicating that the correlation value is statistically insignificant, despite the assumption that the correlation between UIFIN and UIGER is strong at 0.475.

Figure 3.2. Finland – Google Trends based Uncertainty Index of Finland and Germany

Table 3.3.1. Finland – Correlation between Finland's and Germany's Uncertainty Indices

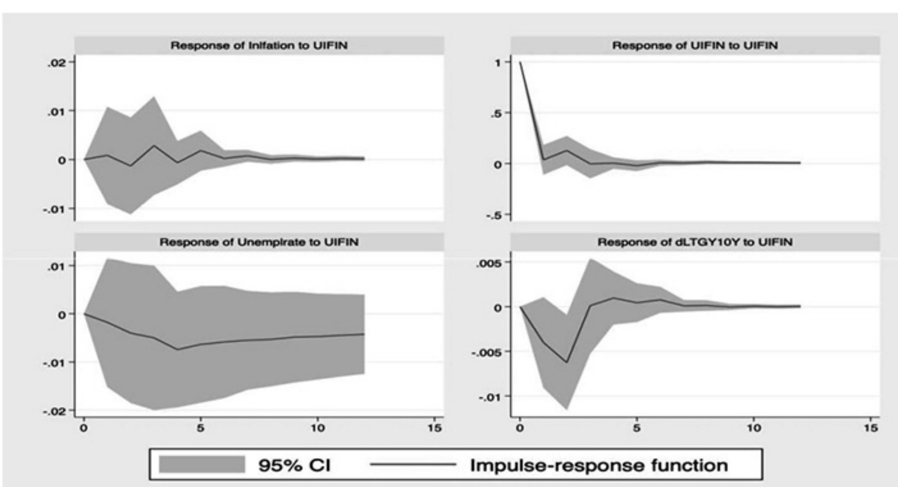
	UI GER	UI FIN
UI GER	1	
UI FIN	0.47522197	1

Table 3.3.2. Finland – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
UI FIN/UI GER	7.205929379	1

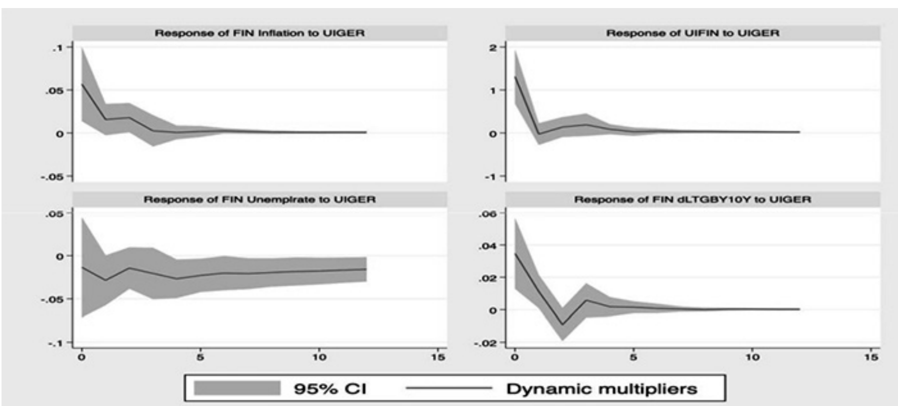
We created a new variable called dLTGBY10Y using the variable's initial differences. The four variables' reactions to a shock with a UIFIN one standard deviation are depicted in the following IRF graphs. Notably, dLTGBY10Y and Unemplrate both responded negatively. On the other hand, the latter reacts with a much smaller magnitude and, by the fourth period, returns and hovers around zero. Unlike Unemplrate, which shows a negative value throughout the course of the twelve periods. However, following the fifth period, the UIFIN response to UIFIN rapidly fades away.

Figure 3.3.1. Finland– Impulse Response Functions to a UIFIN shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



The dynamic-multiplier function used to calculate the effects of a unit increase in an exogenous variable on the endogenous variables over time is the subject of the following set of IRF graphs. The exogenous variable is the German Uncertainty Index. It is evident that the Unemplrate in this instance stayed below zero for the entire duration. In contrast, dLTGBY10Y only twice recorded a negative value. Conversely, FIN Inflation showed positive numbers, reaching a maximum of 0.017% during the second period. Last but not least, UIFIN's response to UIGEIR increased to 0.19% after two periods, having dropped below zero in the first.

Figure 3.3.2. Finland – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



3.4. France

The economic indicators and our Uncertainty Index for France are shown in the following figure. It's also critical to note that uncertainty in the nation appears to have been heightened by the COVID-19 epidemic crisis. The unemployment rate and UIFRA have a strong correlation (-0.463), but the correlation between LTGBY10Y and the unemployment rate is very weak (-0.006).

Figure 4.1. France – Depiction of the Uncertainty Index (UI) and inflation index from January 2008 till December 2022

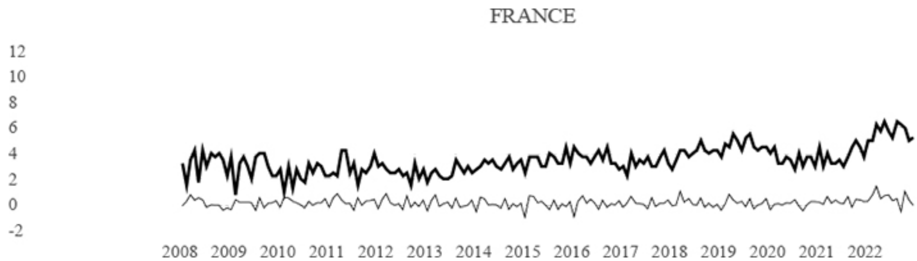


Table 4.1. France – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI France
Unemployment rate	1			
LTGBY 10Y	-0.006390864	1		
Inflation	-0.186177492	0.013406709	1	
UI France	-0.463433287	-0.379678472	0.169435258	1

According to data from Baker et al. (2016), France is one of the Eurozone countries included in the Economic Policy Uncertainty index (EPU). France's CCI data was also available. The positive correlation between CCI FRA and UI FRA is not desirable, whereas the correlation between EPU FRA and UI FRA, at 0.222, is. However, Table 4.2.1 suggests that the correlation between the *t*-statistic and *p*-value is statistically significant for EPU FRA/UI FRA and statistically insignificant for CCI FRA/UI FRA.

Table 4.2.1. Finland – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI FIN), Economic Policy Uncertainty index (EPU) and the Consumer Confidence Index (CCI FIN)

	UI FRA	CCI FRA	EPU FRA
UI FRA	1		
CCI FRA	0.10285576	1	
EPU FRA	0.22254198	0.05580796	1

Table 4.2.2. France – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
CCI/UI	1.379583955	0.169445408
EPU/UI	3.045450743	0.002676319
CCI/EPU	0.74573333	0.456812169

The Germany and France Uncertainty Indices are shown in the following figure. With the two exceptions in 2010–2011 and 2022—two of the largest economies in Europe—the two indices have followed the same trajectory for the entire fifteen years. At the 1% confidence level, the correlation of 0.538 is statistically significant.

Figure 4.2. France – Google Trends based Uncertainty Index of France and Germany

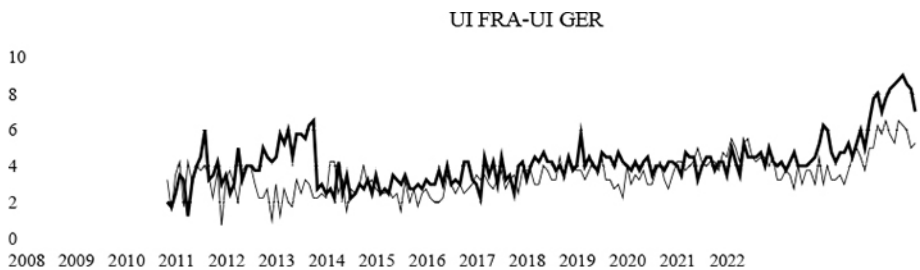


Table 4.3.1. France – Correlation between France's and Germany's Uncertainty Indices

	UI GER	UI FRA
UI GER	1	
UI FRA	0.53805842	1

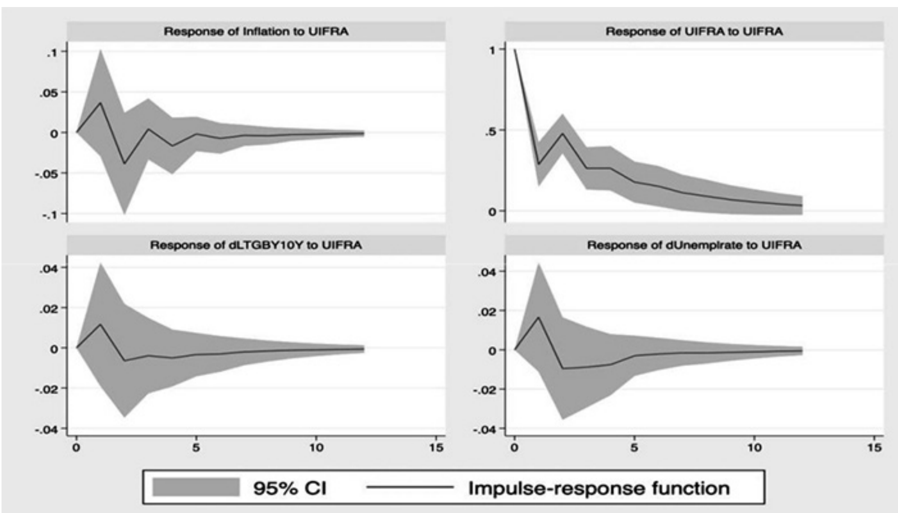
Table 4.3.2. France – *t*-statistic and p-value prices from correlations

	t-statistic	*p*-value
UI FRA/UI GER	8.516467387	6.73576E-15

The first set of IRF graphs shows how the variables LTGBY10Y, UIFRA, unemployment rate, and inflation react to a shock of one standard deviation. It has been noted that dLTGBY10Y and dUnemplrate have a tendency to move in tandem with a UIFRA shock.

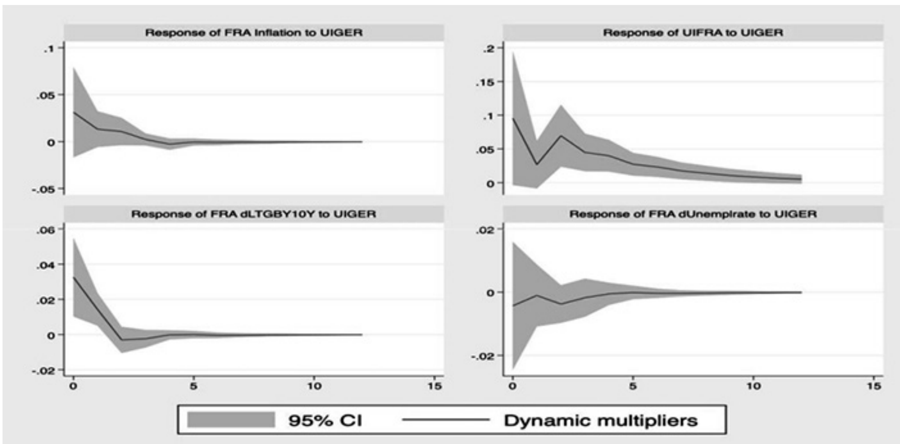
The first few periods see a slight increase in inflation, which quickly turns negative after the second period and stays below zero for the remaining periods. Unlike the other European countries examined thus far, which show a significant decline shortly after the first or second period, UIFRA appears to insist on remaining relatively high for the first five periods following a UIFRA shock.

Figure 4.3.1. France– Impulse Response Functions to a UIFRA shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



The dynamic-multiplier function used to calculate the effects of a unit increase in an exogenous variable on the endogenous variables over time is displayed in the second IRF graphs. The exogenous variable is the German Uncertainty Index. For every relevant period, FRA dUnemplrate was able to move below zero. The dLTGBY10Y did the same, recording only positive values for the first two periods. It is noteworthy that UIGER's impact is now lower than that of another uncertainty index, in our case UIFRA, for the first time. UIFRA's response to UIGER is nearly ten times smaller than UIFRA's response to it.

Figure 4.3.2. France – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



3.5. Germany

The following figure shows the Uncertainty Index and Germany's economic indicators. Beginning in 2008, the unemployment rate reached its highest point of 8%. Despite the COVID-19 pandemic and the Russian invasion of Ukraine, the rate continued to decline, ending in 2022 at slightly over 3%. Given that there were a few precariousness periods in the time horizon under examination and that some people would have predicted the exact opposite result, this is an intriguing fact. The variables' correlations support the economic theory. The unemployment rate and the 10-year government bond yield show a very strong correlation of 0.894, while the unemployment rate and inflation show a negative correlation of -0.121. As was mentioned at the outset of the study, rising unemployment rates typically result in lower inflation and higher spreads on high-yield bonds.

Figure 5.1. Germany – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

Table 5.1. Germany – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Germany
Unemployment rate	1			
LTGBY 10Y	0.894308655	1		
Inflation	-0.121393722	-0.04426441	1	
UI Germany	-0.383694292	-0.217736884	0.25633993	1

Both of Germany's current uncertainty indices were accessible online. As anticipated, the correlations are strong and positive for EPU GER/UIGER and negative and moderate for CCI GER/UI GER. Table 5.2.2 displays that all correlation values are statistically significant, with p-values significantly below 0.01.

Table 5.2.1. Germany – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI GER), the Consumer Confidence Index (CCI GER) and the Economic Policy Uncertainty index (EPU GER)

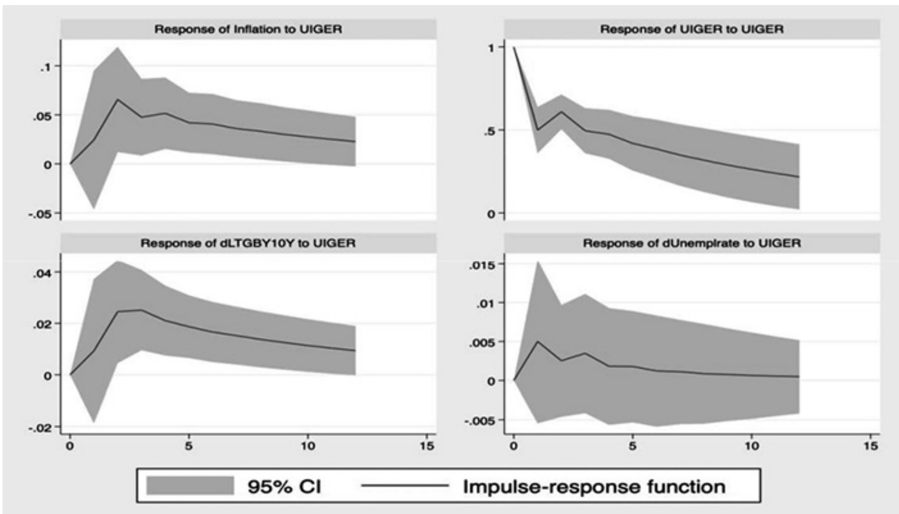
	UI GER	CCI GER	EPU GER
UI GER	1		
CCI GER	-0.3808019	1	
EPU GER	0.68906979	-0.5001428	1

Table 5.2.2. Germany – t-statistic and p-value prices from correlations

	t-statistic	p-value
CCI/UI	-5.4945047	1.33883E-07
EPU/UI	12.68580277	1.10373E-26
CCI/EPU	-7.705747471	8.77322E-13

As Germany is the exogenous variable used in the Dynamic-Multiplier Functions, as previously mentioned, we only have one set of IRF graphs for the country's impulse on the Uncertainty Index to the economic indicators and the Uncertainty Index itself. UIGER's reaction to a one-standard deviation of UIGER is crucial since it requires to remain significantly above 0.20% throughout the analysis period. The UIGER shock has had a positive impact on all economic indicators, with the exception of dUnemplrate, whose response has been less significant.

Figure 6.3.1. Germany– Impulse Response Functions to a UIGER shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



3.6. Greece

Regarding Greece's economic indicators, it is of concern the levels of unemployment rate the country reached by 2013. It took nearly ten years for the reported 26% unemployment rate to drop. It's also critical to recognize that, in contrast to the COVID-19 pandemic crisis, the 2008 financial crisis had a profound impact on the nation's level of uncertainty. This statistic may indicate that public trust in the government has returned following years of mistrust. When the right signals are present, the correlations between inflation and the 10-year government bond yield appear to follow theory.

Figure 6.1. Greece – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

Table 6.1. Greece – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Greece
Unemployment rate	1			
LTGBY 10Y	0.420297043	1		
Inflation	-0.1250371	-0.052882408	1	
UI Greece	-0.326040337	-0.17805485	-0.005446191	1

Below is a correlation between our Uncertainty Index and the current uncertainty indices. Greece possessed data pertaining to the EPU and CCI indices. The correlations' t-statistics and p-values are shown in Table 12.2.2. Statistically speaking, the CCI/UI is more significant than the EPU/UI. Even though the correlation is only 0.152, it is still valid. Even though a positive correlation defies economic theory, it may mean that people will save more and consume less because, in certain economies, the insurance industry offers no security, which encourages people to keep consuming.

Table 6.2.1. Greece – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI GRE), the Consumer Confidence Index (CCI GRE) and the Economic Policy Uncertainty index (EPU GRE)

	UI GRC	CCI GRC	EPU GRC
UI GRC	1		
CCI GRC	0.15205435	1	
EPU GRC	-0.0903901	-0.2117684	1

Table 6.2.2. Greece – t-statistic and p-value prices from correlations

	t-statistic	p-value
CCI/UI	2.052524498	0.041582465
EPU/UI	-1.210911222	0.227534737
CCI/EPU	-2.89090931	0.004319737

Greece's Uncertainty Index seems to be moving in a similar direction as Germany's Uncertainty Index. Greece's UI paradoxically fluctuates very close to Germany's UI for the remaining years of the analysis, despite the first three years of analysis. At the 1% confidence level, the 0.258 correlation is statistically significant but is regarded as weak.

Figure 6.2. Greece – Google Trends based Uncertainty Index of Greece and Germany

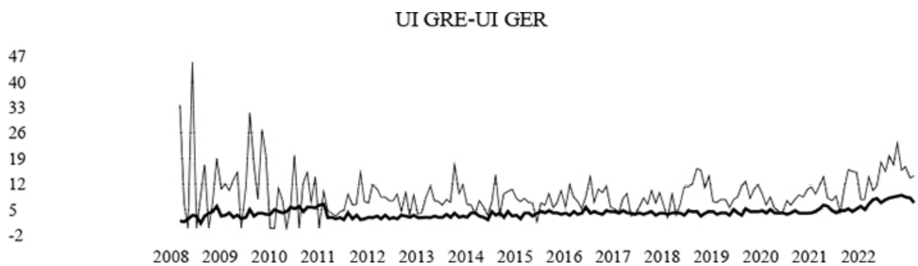


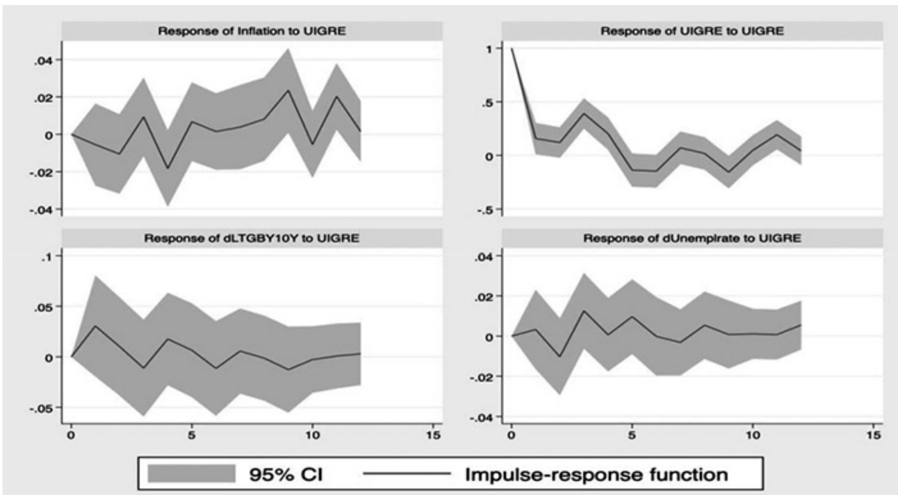
Table 6.3.1. Greece – Correlation between Greece's and Germany's Uncertainty Indices

	UI GER	UI GRC
UI GER	1	
UI GRC	0.25806665	1

Table 6.2.2. Greece– t-statistic and p-value prices from correlations

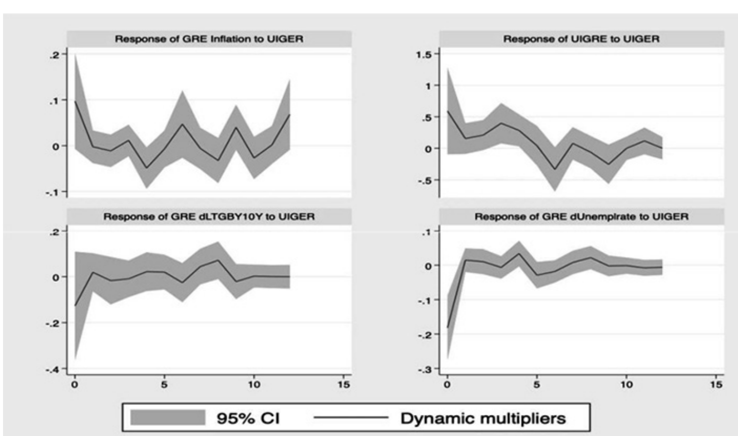
	t-statistic	p-value
UI GRC/UI GER	3.56375314	0.000469391

Figure 6.3.1. Greece– Impulse Response Functions to a UIGRE shock. Sample: 2008M1 – 2022M12. VAR (9) estimated with an exogenous variable (UIGER). 95% confidence interval



The dynamic-multiplier function used to calculate the effects of a unit increase in an exogenous variable on the endogenous variables over time is depicted in the following IRF graphs. The exogenous variable is the German Uncertainty Index. For nearly the entire twelve periods, the variables appear to oscillate around zero. Once more, Greece is the first nation whose UI gradually drops below zero following a one unit increase in UIGER.

Figure 6.3.2. Greece – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



3.7. Ireland

The created Uncertainty Index and economic indicators are shown for Ireland in the following figure. Ireland's unemployment rate peaked in 2010 at about 15%, and it didn't start to decline until after 2013, when it eventually returned to levels it was in 2008 by 2018. The correlations' signs seem to support the economic theory.

Figure 7.1. Ireland – Depiction of the Uncertainty Index (UI) and other economic indices from January 2008 till December 2022

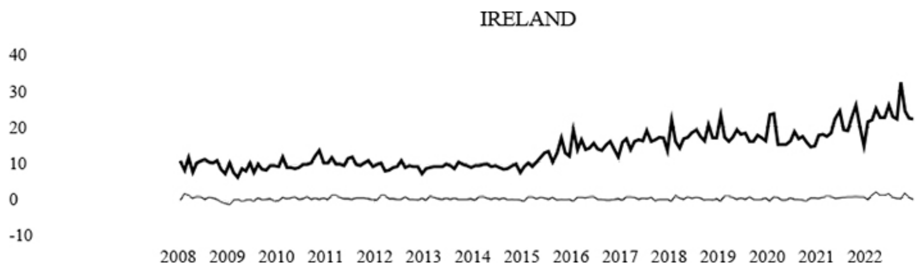


Table 7.1. Ireland – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Ireland
Unemployment rate	1			
LTGBY 10Y	0.786847454	1		
Inflation	-0.149745054	-0.04827397	1	
UI Ireland	-0.763712695	-0.598559992	0.235645325	1

According to Baker et al. (2016), data on Economic Policy Uncertainty (EPU) are available for a number of Eurozone nations, including Ireland. At 0.397, the correlation between UI IRL and CCI IRL is statistically significant and positive. Once more, it is positive and statistically significant between EPU IRL and UI IRL, at 0.508. As was already mentioned, the ideal sign for CCI/UI is negative; however, certain nations exhibit a positive sign, possibly as a result of people's preference for consumption over saving money and the lack of significant security in the insurance industry.

Table 7.2.1. Ireland – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI IRL), the Consumer Confidence Index (CCI IRL) and the Economic Policy Uncertainty index (EPU IRL)

	UI IRL	CCI IRL	EPU IRL
UI IRL	1		
CCI IRL	0.39730098	1	
EPU IRL	0.5089474	-0.035703	1

Table 7.2.2. Ireland – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
CCI/UI	5.776096057	3.3474E-08
EPU/UI	7.888271462	2.98666E-13
CCI/EPU	-0.476641189	0.634202247

Germany's and Ireland's Uncertainty Indices are displayed on the graph. Throughout the entire analysis period, the UIIRL was higher than the UIGER. UIIRL nearly doubled in value just after 2015, and by the end of 2022, it had tripled its 2008 level price. Germany's Uncertainty Index doubled only briefly between 2010 and 2022, in contrast to Ireland. The following table's *p*-value indicates that the correlation is statistically significant at 0.651.

Figure 7.2. Ireland – Google Trends based Uncertainty Index of Ireland and Germany

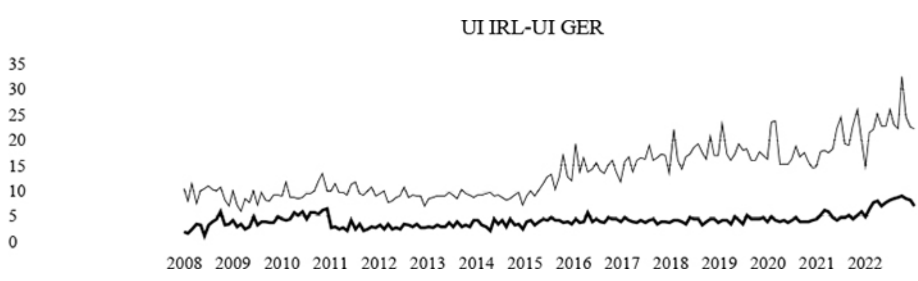


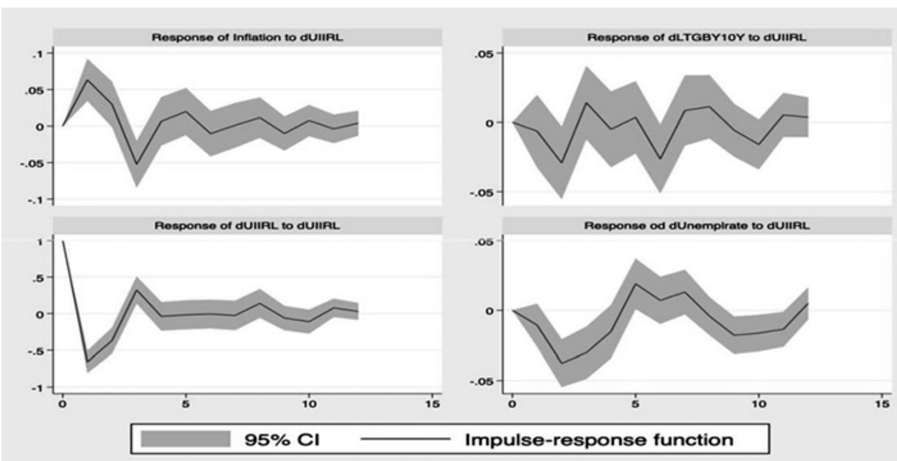
Table 7.3.1. Ireland – Correlation between Ireland's and Germany's Uncertainty Indices

	UI GER	UI IRL
UI GER	1	
UI IRL	0.65194655	1

Table 7.3.2. Ireland – t-statistic and p-value prices from correlations

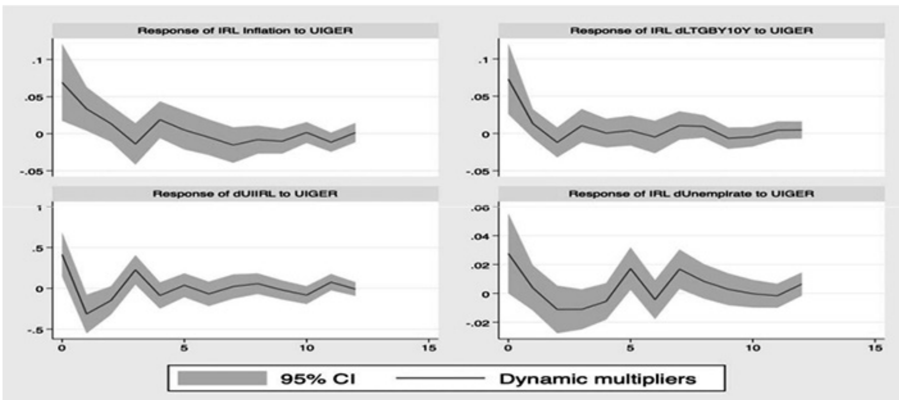
	t-statistic	p-value
UI IRL/UI GER	11.47099328	3.68837E-23

Figure 7.3.1. Ireland– Impulse Response Functions to a UIIRL shock. Sample: 2008M1 – 2022M12. VAR (7) estimated with an exogenous variable (UIGER). 95% confidence interval



About the second IRF graphs, they show how a unit increase in an exogenous variable affects the endogenous variables over time using a dynamic-multiplier function. The exogenous variable is the German Uncertainty Index. All variables produced values that were equally positive and negative, with the exception of dLTGBY10Y, for which the first, third, and seventh periods had higher levels of positive values.

Figure 7.3.2. Ireland – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (7) estimated with an exogenous variable (UIGER). 95% confidence interval



3.8. Italy

Italy is one of the few countries that recorded an Uncertainty Index level less than that of their unemployment rate. Italy's unemployment rate peaked in 2014 at 13% and briefly recovered to 2008 levels in 2020 and 2022. Unlike other Eurozone nations, the Uncertainty Index did not show sharp increases; instead, it varied between the 0 and 9 price levels. The Table displays the relationships between the Uncertainty Index and the economic indicators.

Figure 8.1. Italy – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

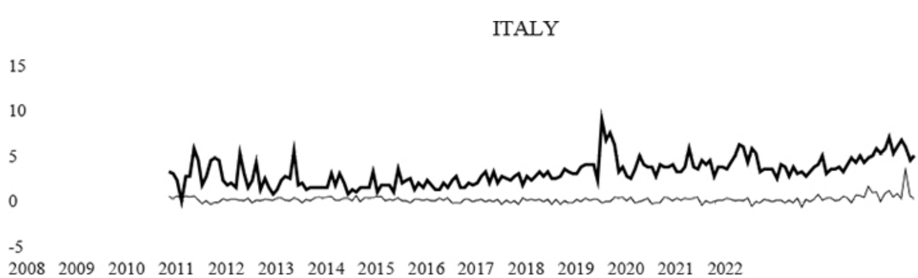


Table 8.1. Italy – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long- term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Italy
Unemployment rate	1			
LTGBY 10Y	-0.302728159	1		
Inflation	-0.279655512	0.11117023	1	
UI Italy	-0.10675958	-0.440526252	0.155465839	1

For Italy there is data for the EPU and CCI uncertainty indices. There is a statistically significant correlation between CCI ITA and UI ITA, but there is a statistically insignificant correlation between EPU ITA and UI ITA, according to the results in Table 8.

Table 8.2.1. Italy – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI ITA), the Consumer Confidence Index (CCI ITA) and the Economic Policy Uncertainty index (EPU ITA)

	UI ITA	CCI ITA	EPU ITA
UI ITA	1		
CCI ITA	0.30312711	1	
EPU ITA	-0.0917678	-0.3688628	1

Table 8.2.2. Italy – t-statistic and p-value prices from correlations

	t-statistic	p-value
CCI/UI	4.243894768	3.52972E-05
EPU/UI	-1.229522926	0.220498535
CCI/EPU	-5.294598958	3.48443E-07

Again, Italy is among the few countries whose uncertainty index was able to nearly exactly match that of Germany. One possible explanation for the slight increase in UI ITA in 2017 could be the nation's series of earthquakes that year. At the 1% confidence level, the variables' 0.517 correlation is thought to be strong and statistically significant.

Figure 8.2. Italy – Google Trends based Uncertainty Index of Italy and Germany

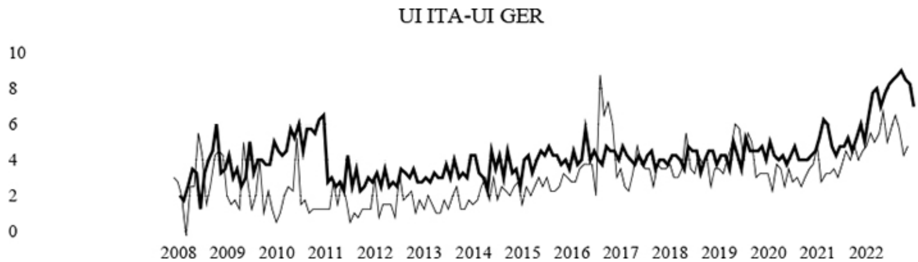


Table 8.3.1. Italy – Correlation between Italy's and Germany's Uncertainty Indices

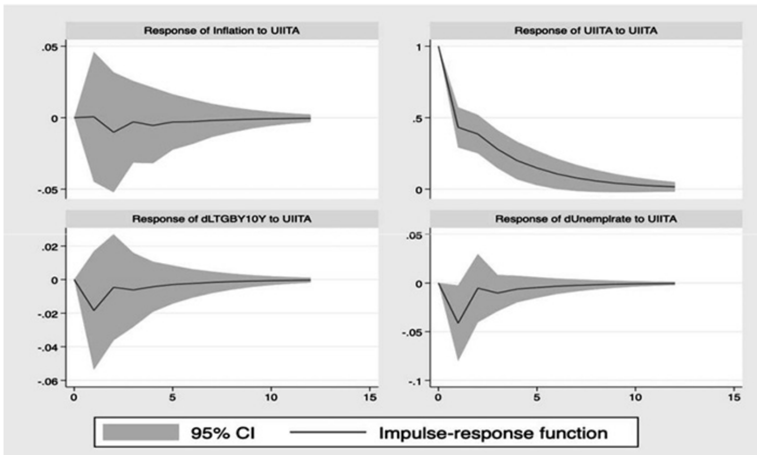
	UI GER	UI ITA
UI GER	1	
UI ITA	0.51783404	1

Table 8.3.2. Italy – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
UI ITA/UI GER	8.075886092	9.75054E-14

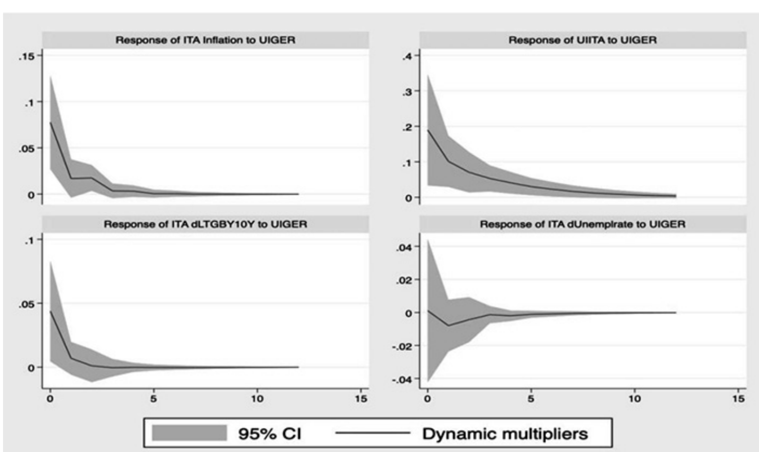
The variables' responses to a single UIITA standard deviation are displayed in the IRF graphs. It is evident that over the course of the year, negative values dominate for dLTG-BY10Y, dUnemplrate, and inflation. As opposed to UIITA, which appeared to maintain its high price during the first five months.

Figure 8.3.1. Italy– Impulse Response Functions to a UIITA shock. Sample: 2008M1 – 2022M12. VAR (7) estimated with an exogenous variable (UIGER). 95% confidence interval



The second set of IRF graphs shows how a dynamic multiplier function affects the endogenous variables over time in response to a unit increase in an exogenous variable. The exogenous variable is the German Uncertainty Index. With the exception of ITA dUnemplrate, all variables start out positively and appear to respond to UIGER shocks by creating a smooth downward slope that eventually converges to zero by the fifth period. Throughout the periods under examination, UIITA was able to stay below zero, with the fourth period seeing the lowest value at -0.002%.

Figure 8.3.2. Italy – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (7) estimated with an exogenous variable (UIGER). 95% confidence interval



3. 9. Netherlands

The Uncertainty Index and the economic indicators for the Netherlands are shown in Figure 9.1. The unemployment rate in the Netherlands peaked in 2014 at 9%. With the exception of the Uncertainty Index sign, the correlation signals appear to be consistent with theory. 2014 saw a modest increase in the Uncertainty Index, tripling its value from the previous year. After that, the index never went back to its 2013 levels; instead, it grew over time, reaching its all-time high of 2008 by the end of 2022.

Figure 9.1. Netherlands – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2020

Table 9.1. Netherlands – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Netherlands
Unemployment rate	1			
LTGBY 10Y	0.099627988	1		
Inflation	-0.181130983	-0.037533865	1	
UI Netherlands	-0.499545797	-0.442901685	0.204133209	1

For Netherlands there is data both for the CCI and EPU indices. Nevertheless, the data was only accessible through December 2020. As a result, the 2008–2020 timeframe is covered in the correlation analysis between the indices that follows. Since one would anticipate the opposite signs for each variable, the correlations' results were not desirable. As shown in Table 9.2.2, they were discovered to be statistically significant nonetheless.

Table 9.2.1. Netherlands – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI NLD), the Consumer Confidence Index (CCI NLD) and the Economic Policy Uncertainty index (EPU NLD)

	UI NLD	CCI NLD	EPU NLD
UI NLD	1		
CCI NLD	0.29780402	1	
EPU NLD	-0.180406	-0.6367986	1

Table 9.2.2. Netherlands – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
CCI/UI	4.162045502	4.90444E-05
EPU/UI	-2.447067666	0.015373568
CCI/EPU	-11.01896298	7.36809E-22

As previously mentioned, the global financial crisis causes the Netherlands' Uncertainty Index to spike in 2008, fall back in 2010, rise slightly in 2014, and stay there for the remaining six years, until December 2020. The Uncertainty Indices of the two nations have a statistically significant correlation of 0.338.

Figure 9.2. Netherlands – Google Trends based Uncertainty Index of Netherlands and Germany

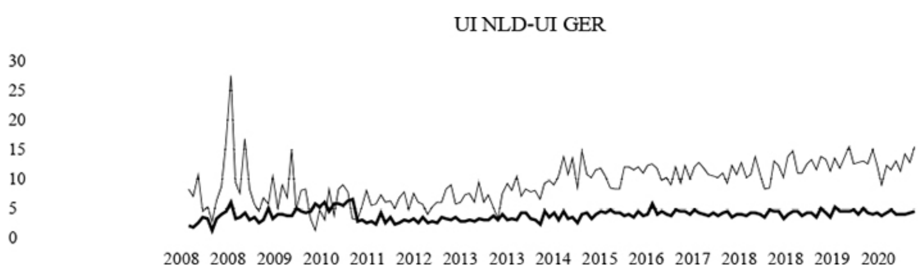


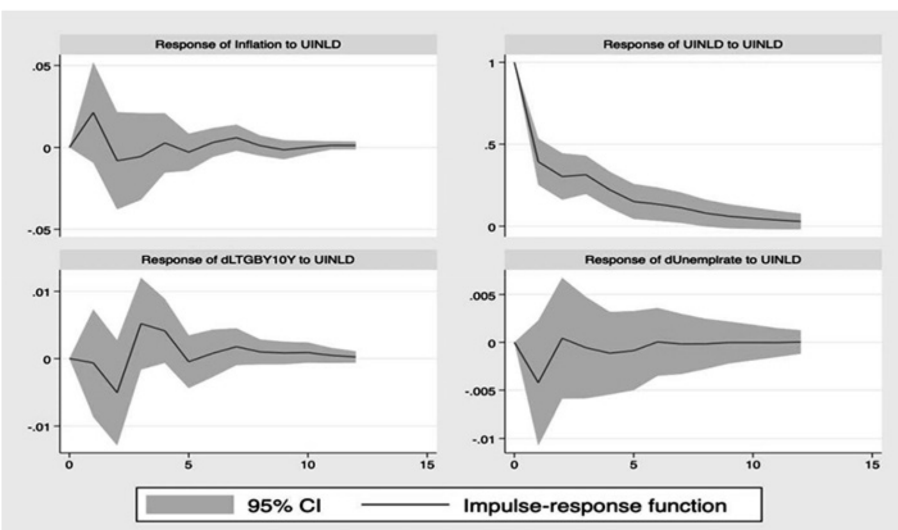
Table 9.3.1. Netherlands – Correlation between Netherlands's and Germany's Uncertainty Indices

	UI GER	UI NLD
UI GER	1	
UI NLD	0.33847303	1

Table 9.3.2. Netherlands – *t*-statistic and *p*-value prices from correlations

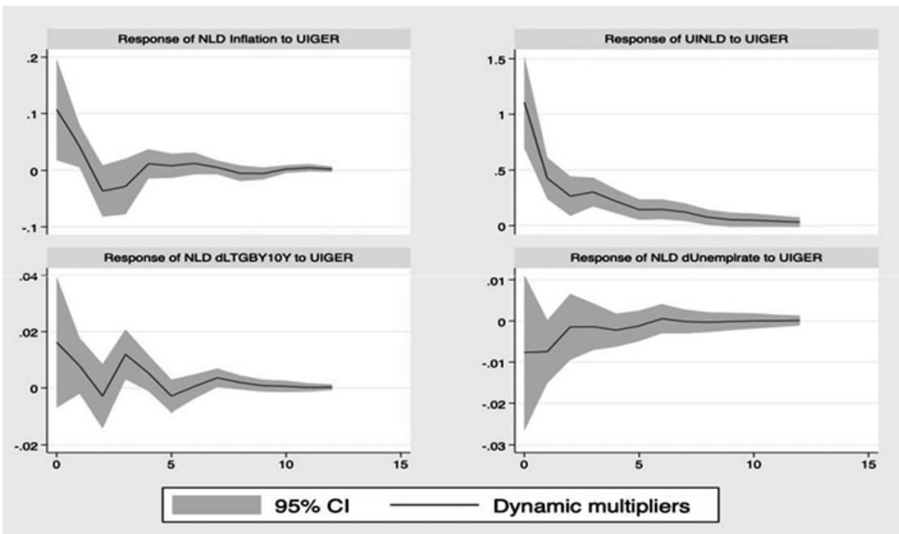
	t-statistic	*p*-value
UI NDL/UI GER	4.799052234	3.36197E-06

Figure 9.3.1. Netherlands– Impulse Response Functions to a UINLD shock. Sample: 2008M1 – 2022M12. VAR (3) estimated with an exogenous variable (UIGER). 95% confidence interval



A dynamic-multiplier function, shown in Figure 9.3.2, is used to calculate the time-dependent effect of a unit increase in an exogenous variable on the endogenous variables. The exogenous variable is the German Uncertainty Index.

Figure 9.3.2. Netherlands – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (3) estimated with an exogenous variable (UIGER). 95% confidence interval



3.10. Portugal

The economic indices and the Uncertainty Index for Portugal are displayed below. Portugal's unemployment rate peaked in 2013 at 19%, but it then steadily declined until shortly after 2017 when it finally reached the 2008 level. Ironically, during the 2008 financial crisis, our Uncertainty Index shows zero values. All of the variables' correlations are shown in Table 10.1.

Figure 10.1. Portugal – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

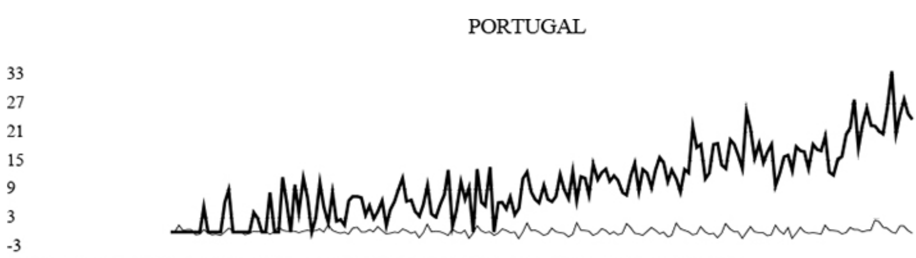


Table 10.1. Portugal – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Portugal
Unemployment rate	1			
LTGBY 10Y	0.730216599	1		
Inflation	-0.085348002	0.011189386	1	
UI Portugal	-0.630545575	-0.557599162	0.132669325	1

We did not find data for the Economic Policy Uncertainty Index; instead, we only found data for the Consumers Confidence Index. At the 1% confidence level, the correlation that was provided was equal to 0.295 and was statistically significant

Table 10.2.1. Portugal – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI PRT) and the Consumer Confidence Index (CCI PRT)

	UI PRT	CCI PRT
UI PRT	1	
CCI PRT	0.29595466	1

Table 10.2.2. Portugal – t-statistic and p-value prices from correlations

	t-statistic	p-value
CCI/UI	4.133709677	5.48996E-05

The following figure illustrates how Germany's Uncertainty Index has remained relatively stable over the years, even with a few global crises. However, shortly after 2017, Portugal's Uncertainty Index started to rise. A portion of the unpredictability may be attributed to the four initial deadly wildfires that broke out in central Portugal in June 2017, resulting in numerous fatalities and injuries. The Russian invasion of Ukraine and the COVID-19 epidemic crisis prevented the UI PRT from ever reaching its 2015–2016 levels. With a correlation of 0.604, the relationship between UI PRT and UI GER is regarded as statistically significant and strong.

Figure 10.2. Portugal – Google Trends based Uncertainty Index of Portugal and Germany

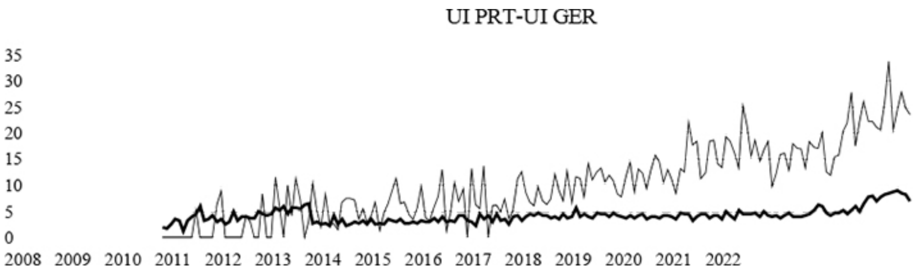


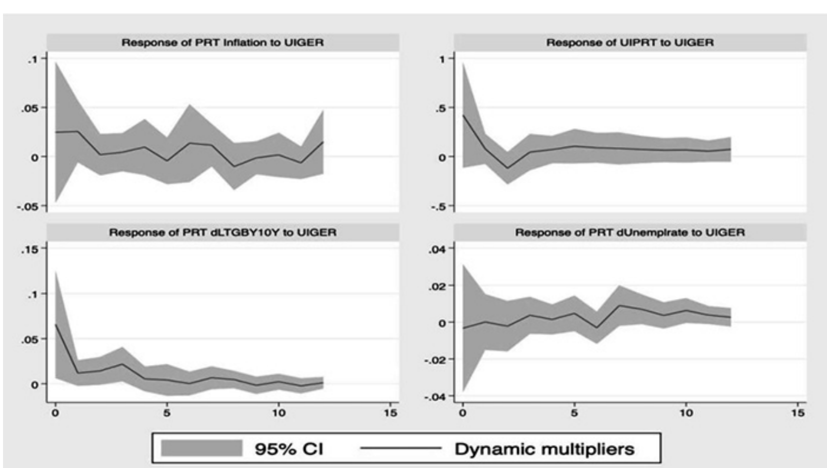
Table 10.3.1. Portugal – Correlation between Portugal's and Germany's Uncertainty Indices

	UI GER	UI PRT
UI GER	1	
UI PRT	0.6047231	1

Table 10.3.2. Portugal – *t*-statistic and *p*-value prices from correlations

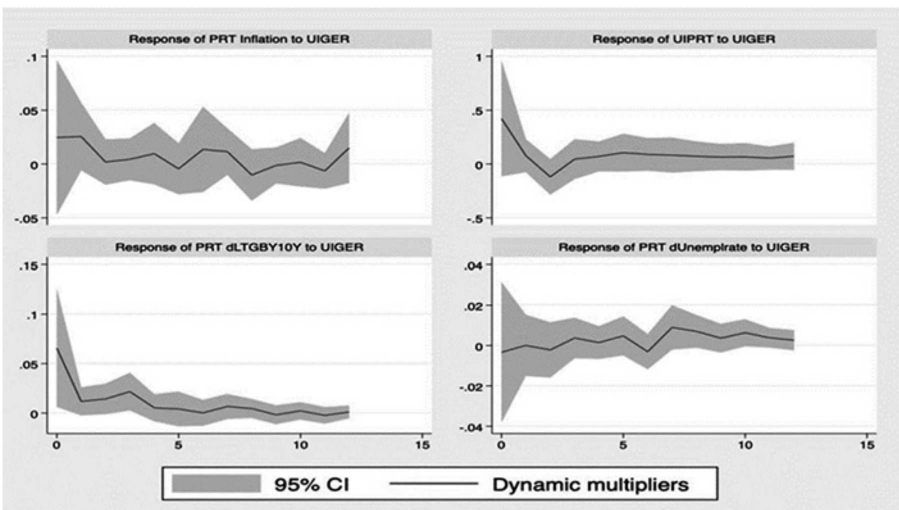
	t-statistic	*p*-value
UI PRT/UI GER	10.13014793	2.48491E-19

Figure 10.3.1. Portugal– Impulse Response Functions to a UIPRT shock. Sample: 2008M1 – 2022M12. VAR (7) estimated with an exogenous variable (UIGER). 95% confidence interval



A dynamic-multiplier function, as shown in Figure 10.3.2 of Portugal, is used to calculate the time-dependent effect of a unit increase in an exogenous variable on the endogenous variables. The exogenous variable is the German Uncertainty Index. Once more, it is seen that the variables are moving very near to zero over the course of the twelve periods, indicating that UIGER's impact on the Portuguese economy is not particularly noteworthy.

Figure 10.3.2. Portugal – Dynamic-Multiplier Functions of a UIGER shock. Sample: *2008M1 – 2022M12*. VAR (7) estimated with an exogenous variable (UIGER). 95% confidence interval



3.11. Spain

The following figure shows the Uncertainty Index and economic indicators for Spain. It is noteworthy that Spain became the second country to record an unemployment rate above 20% in the fifteen years of analysis in 2013, when it reached 26%, a level only Greece attained at roughly the same time. However, despite a high unemployment rate, the Uncertainty Index was relatively low during the first ten years. Following the COVID-19 epidemic crisis, the index slightly increased and remained there until the end of 2022. The correlations between the variables are shown in Table 11.1.1.

Figure 11.1. Spain – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

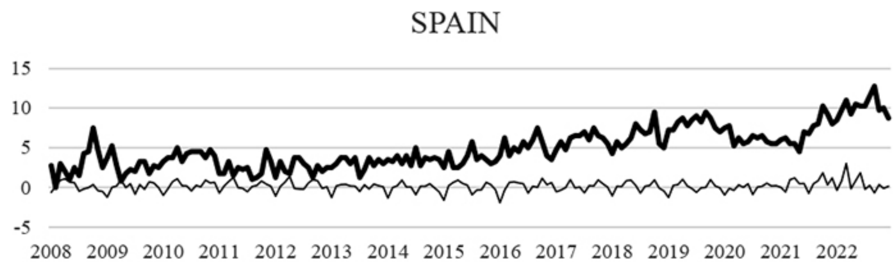


Table 11.1. Spain – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Spain
Unemployment rate	1			
LTGBY 10Y	0.480545888	1		
Inflation	-0.097831779	-0.027478227	1	
UI Spain	-0.579663791	-0.636917648	0.11691803	1

Data for the CCI and EPU uncertainty indices were available for Spain. The correlations between the built Uncertainty Index and the current uncertainty indices were both found to be statistically significant at the 1% confidence level, which means that the correlations are accurate to their respective values. Still, we are confronted with the unwanted consequence of the positive correlation between CCI and UI. Once more, this may be because people are not feeling safe to invest into deposits and instead, they consume their income.

Table 11.2.1. Spain – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI ESP), the Consumer Confidence Index (CCI ESP) and the Economic Policy Uncertainty index (EPU ESP)

	UI ESP	CCI ESP	EPU ESP
UI ESP	1		
CCI ESP	0.299168	1	
EPU ESP	0.28150059	0.01256303	1

Table 11.2.2. Spain – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
CCI/UI	4.182977347	4.51078E-05
EPU/UI	3.913962762	0.000129126
CCI/EPU	0.167624993	0.867068599

Spain's Uncertainty Index seems to take the same path as Germany's Uncertainty Index. Spain is the nation with the highest correlation between its Uncertainty Index and Germany's, with a statistically significant correlation as high as 0.720.

Figure 11.2. Spain – Google Trends based Uncertainty Index of Spain and Germany

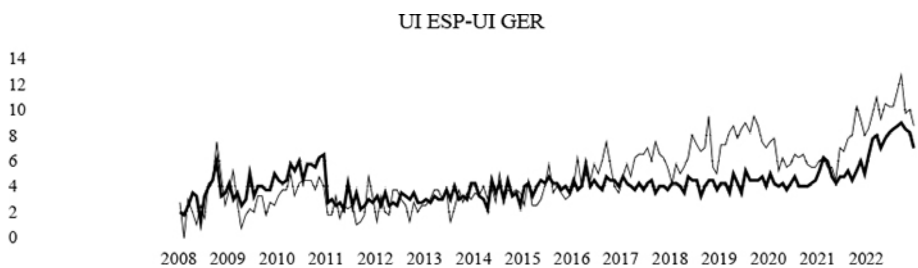


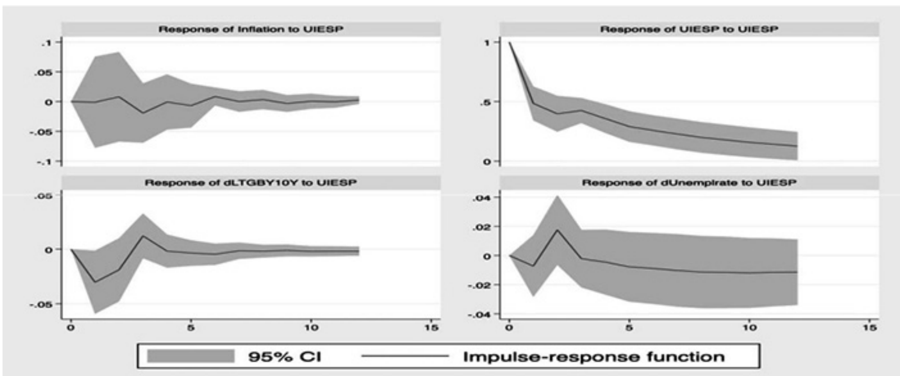
Table 11.3.1. Spain – Correlation between Spain's and Germany's Uncertainty Indices

	UI GER	UI ESP
UI GER	1	
UI ESP	0.72010589	1

Table 11.3.2. Spain – *t*-statistic and *p*-value prices from correlations

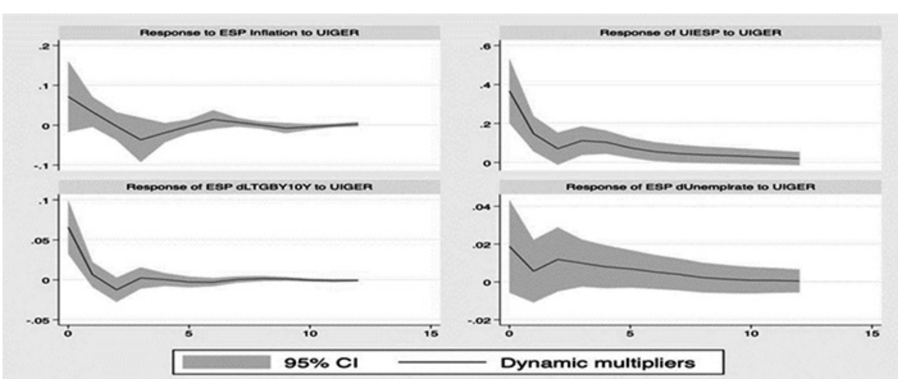
	t-statistic	*p*-value
UI ESP/UI GER	13.84624115	4.60646E-30

Figure 11.3.1. Spain – Impulse Response Functions to a UISVN shock. Sample: 2008M1 – 2022M12. VAR (3) estimated with an exogenous variable (UIGER). 95% confidence interval



A dynamic multiplier function, shown in Figure 11.3.2, is used to calculate the time-dependent effect of a unit increase in an exogenous variable on the endogenous variables. The exogenous variable is the German Uncertainty Index. It appears that every variable travels in the same direction. UIESP and ESP dUnemplrate are the two that are able to remain positive over the periods, while the other two only briefly fell below zero.

Figure 11.3.2. Spain – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (3) estimated with an exogenous variable (UIGER). 95% confidence interval



3.12. LUXEMBOURG

Figure 12.1 shows the inflation rate, unemployment rate, yield on 10-year government bonds, and uncertainty index for Luxembourg. The first nation in Europe to record a positive correlation, at 0.048, between the unemployment rate and the uncertainty index was Luxembourg.

Figure 12.1. Luxembourg – Depiction of the Uncertainty Index (UI) and the inflation index from January 2008 till December 2022

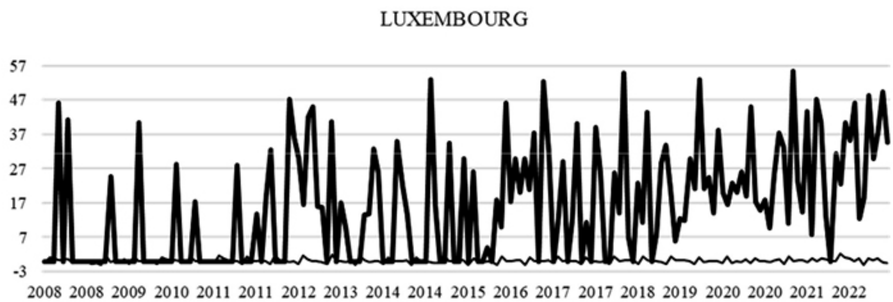


Table 12.1. Luxembourg – Correlation between the Uncertainty Index (UI), unemployment rate, inflation and long-term government bond 10-year yield

	Unemployment rate	LTGBY 10Y	Inflation	UI Luxembourg
Unemployment rate	1			
LTGBY 10Y	-0.603397165	1		
Inflation	-0.129985254	0.003647903	1	
UI Luxembourg	0.048737805	-0.30868969	-0.001107036	1

There were no online data available for Luxembourg's Economic Policy Uncertainty index. The Luxembourg Consumer Confidence Index and the Uncertainty Index we constructed using Google Trends had a negative correlation, measuring -0.022. The correlation is statistically insignificant, according to the *t*-statistic and *p*-value results.

Table 12.2.1. Luxembourg – Measures of uncertainty: Correlation between our constructed Uncertainty Index (UI LUX) and the Consumer Confidence Index (CCI LUX)

	UI LUX	CCI LUX
UI LUX	1	
CCI LUX	-0.0228705	1

Table 12.2.2. Luxembourg – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
CCI/UI	-0.305210459	0.760562469

The graph below displays the uncertainty indices for Germany and Luxembourg. It's clear that, in contrast to UI GER, UI LUX reacts much more aggressively. Consecutive, sharp spikes that occur over the course of the research time horizon define UI LUX. At 0.238, the correlation coefficient between the variables is statistically significant but weak.

Figure 12.2. Luxembourg – Google Trends based Uncertainty Index of Luxembourg and Germany

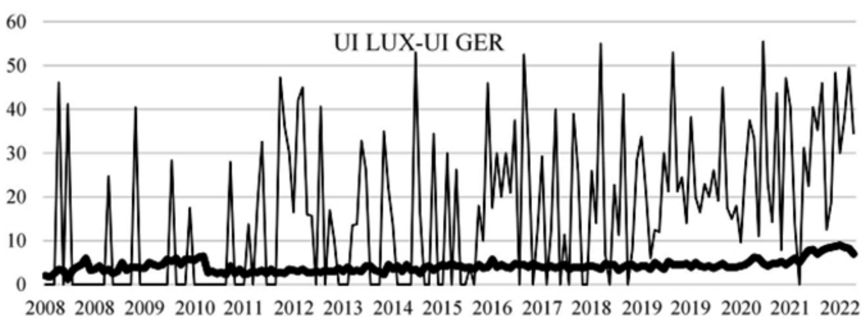


Table 12.3.1. Luxembourg – Correlation between Luxembourg's and Germany's Uncertainty Indices

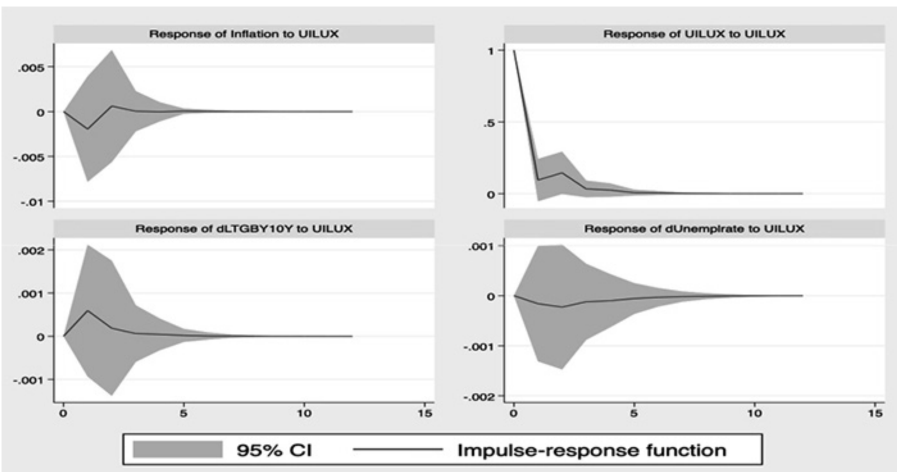
	UI GER	UI LUX
UI GER	1	
UI LUX	0.23839347	1

Table 12.3.2. Luxembourg – *t*-statistic and *p*-value prices from correlations

	t-statistic	*p*-value
UI LUX/UI GER	3.274987841	0.001269911

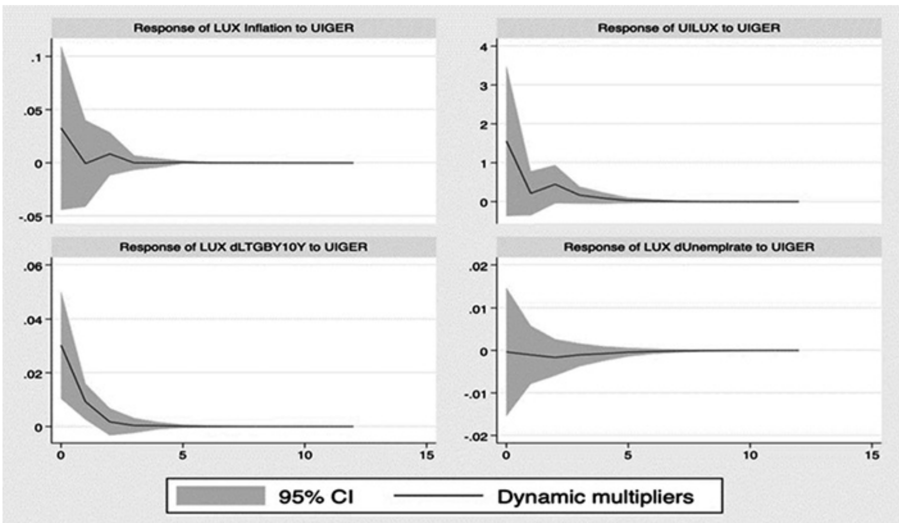
The following IRF graphs display the variables' response to a one-standard deviation of UILUX. It appears that the Luxembourg Uncertainty Index (UI) has very little effect on any economic indicator. Shortly after their initial small response in the first period, the variables converge to zero.

Figure 12.3.1. Luxembourg– Impulse Response Functions to a UILUX shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



A dynamic-multiplier function is used in the following set of IRF graphs to calculate the time-dependent effect of a unit increase in an exogenous variable on the endogenous variables. The exogenous variable is the German Uncertainty Index. UIGER affects Luxembourg's economy more than UILUX does. All response values except for the LUX dUnemplrate ones were recorded as positive.

Figure 12.3.2. Luxembourg – Dynamic-Multiplier Functions of a UIGER shock. Sample: 2008M1 – 2022M12. VAR (2) estimated with an exogenous variable (UIGER). 95% confidence interval



4. **Conclusions**

Mixed results were obtained when a macroeconomic uncertainty index based on Google Trends was constructed. Our Uncertainty Index demonstrated encouraging correlations with other uncertainty indices, including the Economic Policy Uncertainty index (EPU) by Baker et al. (2016) and the Consumer Confidence Index (CCI), for the majority of Eurozone countries, and it was in line with economic theory. In contrast to theoretical expectations, some countries produced unfavorable results when our Uncertainty Index was correlated with the CCI and EPU.

A plausible rationale addressed in the piece concerned the feeble or unstable stability of the nation's insurance industry. In these situations, residents might choose to spend rather than save money even during times of great uncertainty. The narrow scope of the index's construction – only four terms were used, all of which were studied in English without translation into the local tongue – contributed to less desirable and indicative results. When writing about this particular subject, authors frequently concentrate on creating uncertainty indices for one country or, at most, two countries, utilizing forty or more keywords in the process.

Positive results were found for the Impulse-Response and Dynamic-Multiplier Functions. First, the effect of the Uncertainty Index on the economic indicators of each nation was looked at. The responses of each country's variables to a unit increase in the Uncertainty Index of Germany – the biggest economy in Europe – were then examined. When a shock to Germany's uncertainty index occurred, the responses of most European countries' variables appeared more pronounced, explicit, and significant than when a shock occurred to their own uncertainty index. As explored in other papers on Google Trends, if a large number of words were employed in the analysis to create the macroeconomic uncertainty index, the outcomes would probably be more precise, trustworthy, and appropriate for making justified conclusions. For this reason, for future research it would be important the inclusion of more words for the construction of the uncertainty index. further research regarding the topic is suggested.

References

Albert, J. F.; Gómez Fernández, N. (2018) The impact of uncertainty shocks in Spain: SVAR approach with sign restrictions. LSE Research Online Documents on Economics 90402, London School of Economics and Political Science, LSE Library.

Arbatli, E. C.; Davis, S. J.; Ito, A.; Miake, N.; Saito, I. (2017) Policy uncertainty in Japan. Working Paper 23411, National Bureau of Economic Research.

Azqueta-Gavaldón, A.; Hirschbühl, D.; Onorante, L.; Saiz, L. (2023) Sources of economic policy uncertainty in the Euro area. *European Economic Review*, 152, 104373.

Bampinas, G.; Panagiotidis, T.; Rouska, C. (2019) Volatility persistence and asymmetry under the microscope: The role of information demand for gold and oil. *Scottish Journal of Political Economy*, 66(1), 180-197.

Bełej, M. (2022) Does Google Trends show the strength of social interest as a predictor of housing price dynamics? *Sustainability*, 14(9), 5601.

Bilgin, M. H.; Demir, E.; Gozgor, G.; Karabulut, G.; Kaya, H. (2019) A novel index of macroeconomic uncertainty for Turkey based on Google-Trends. *Economics Letters*, 184, 108601.

Bilgin, M. H.; Demir, E.; Gozgor, G.; Karabulut, G.; Kaya, H. (2019) A new method to measure economic and financial uncertainty in Turkey. Available at SSRN: https://ssrn.com/abstract=3337913

Bloom, N. (2009) The impact of uncertainty shocks. *Econometrica*, 77(3), 623-685.

Bonciani, D. (2018) Uncertainty and the Macroeconomy. Doctoral Dissertation, European University Institute, Florence, Italy.

Bontempi, M. E.; Golinelli, R.; Squadrani, M. (2016) A new index of uncertainty based on internet searches: A friend or foe of other indicators? Quaderni-Working Paper DSE N°1062, Univesità di Bologna, Italy.

Bulczak, G. M. (2021) Use of Google Trends to predict the real estate market: Evidence from the United Kingdom. *International Real Estate Review*, 24(4).

Castelnuovo, E.; Tran, T. D. (2017) Google it up! A Google Trends-based uncertainty index for the United States and Australia. *Economics Letters*, 161, 149-153.

Choi, H.; Varian, H. (2012) Predicting the present with Google Trends. *Economic Record*, 88, 2-9.

Dergiades, T.; Milas, C.; Panagiotidis, T. (2015) Tweets, Google Trends, and sovereign spreads in the GIIPS. *Oxford Economic Papers*, 67(2), 406-432.

Donadelli, M. (2015) Google search-based metrics, policy-related uncertainty and macroeconomic conditions. *Applied Economics Letters*, 22(10), 801-807.

Ettredge, M.; Gerdes, J.; Karuga, G. (2005) Using web-based search data to predict macroeconomic statistics. *Communications of the ACM*, 48(11), 87-92.

Hayford, M. D. (2000) Inflation uncertainty, unemployment uncertainty and economic activity. *Journal of Macroeconomics*, 22(2), 315-329.

Kropiński, P.; Anholcer, M. (2022) How Google Trends can improve market predictions: The case of the Warsaw Stock Exchange. *Economics and Business Review*, 8(2), 7-28.

Limnios, A. C.; You, H. (2021) Can Google Trends improve housing market forecasts? *Curiosity: Interdisciplinary Journal of Research and Innovation*, 1(2).

Milas, C.; Panagiotidis, T.; Dergiades, T. (2021) Does it matter where you search? Twitter versus traditional news media. *Journal of Money, Credit and Banking*, 53(7), 1757-1795.

Moore, A. (2017) Measuring economic uncertainty and its effects. *Economic Record*, 93(303), 550-575.

Perju-Mitran, A.; Zirra, D.; Mitran, R. A. (2022) Assessing public interest towards renewable wind energy using Google Trends. *Preprints*, 2022020023. https://doi.org/10.20944/preprints202202.0023.v1

Schütze, F. (2020) Google Trends topic-based uncertainty: A multi-national approach. 3rd International Conference on Advanced Research Methods and Analytics (CARMA2020) Universitat Politècnica de València, València. DOI: http://dx.doi.org/10.4995/CARMA2020.2020.11622

Zayed, B. A., Talaia, A. M., Gaaboobah, M. A., Amer, S. M., & Mansour, F. R. (2023) Google Trends as a predictive tool in the era of COVID-19: A scoping review. *Postgraduate Medical Journal*, Aug 22, 99(1175): 962-975.

Marcelo Santos
Univ Coimbra, CeBER, Faculty of Economics
marcelosantos@fe.uc.pt
orcid.org/0000-0003-2065-6579
Susana Garrido
Univ Coimbra, CeBER, Faculty of Economics
garrido.susana@fe.uc.pt
orcid.org/0000-0001-5229-3130
Tiago Sequeira
Univ Coimbra, CeBER, Faculty of Economics
tiago.n.sequeira@fe.uc.pt
orcid.org/0000-0002-5501-1562

DOI: https://doi.org/10.14195/2183-203X_57_2

Management and Human Capital Employment: An Overlooked Relationship

Capacidades de Gestão e Emprego Qualificado: Uma Relação Pouco Explorada

Marcelo Santos
Susana Garrido
Tiago Sequeira

Received for publication: September 5, 2023
Revision accepted for publication: November 1, 2023

ABSTRACT
In this article we look at data on management and skills demand of firms in existing databases and we highlight the strong positive relationship between both variables. We develop a model that explains this relationship and calibrate it in order to present quantitative results, which we then compare with our own estimates. We discover that a simple model with management as a technology can replicate well the estimated influence of management in the skills that firms require. We also present evidence of the influence of the sub-items of management on skill requirements and found that aside from the talent component of management, target and performance components greatly influence the demand for skills.
Keywords: Management practices; productivity; human capital.

JEL classification: L2; M2; M5; O32; O33; O34.

Aknowledgment: This work has been funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020.

1. Introduction

Differences in management practices (or management quality) has been shown to be an important determinant of differences in firms', industries' and countries' productivity levels: about a quarter of cross-country and within-country TFP gaps can be accounted for by management practices. A review article that summarizes the main results of this recent literature, which began with the article of Bloom and Van Reenen (2007], is Bloom et al. (2014). Management scores are constructed and made publicly available by the World Management Survey (WMS) – initially described in Bloom and Van Reenen (2007] – and have been widely used in this literature. A more recent description is provided in Bloom et al. (2016). The WMS questions address practices that are likely to be associated with delivering existing goods or services more efficiently, focusing on production (lean), human resources management (talent), and management of goals and performance (target and performance, respectively). Managers are the interviewees.

Higher management scores are positively and significantly associated with higher productivity, firm size, profitability, sales growth, market value, and survival. For example, Bloom et al. (2012a) use a database of 10,000 organizations across 20 countries and estimate production functions in which they regress real firm sales on the management score including controls for other inputs (e.g. labor, capital, employee education) and other covariates (e.g. firm age, noise controls, industry, country and year dummies). In the cross section their results show that a one standard deviation increase in management is associated with an increase in TFP of 15%. This relationship is monotonically increasing. The paper also discusses the possibility of nonlinear relationships on the top of the management scores distributions. Meagher and Strachan (2013] apply Bayesian techniques to the Bloom and Van Reenen (2007) data for four countries and also find that there is some convexity for high scores. They interpret this as consistent with the idea that there is complementarity between multiple managerial practices (as in Gibbons and Henderson, 2013); Milgrom and Roberts (1990). Bloom and Van Reenen (2010) discuss why management practices differ across firms and countries. Bloom et al. (2012c) extended the empirical analysis to the transition economies. Competition, multinational and private ownership, and human capital are strongly correlated with better management practices, which means, according to the authors, that more competition, openness, and education in those economies would push management practices upward. Not only manufacturing firms, but also hospitals, schools and retailing sectors have been analyzed (Bloom et al., 2012b; Bloom et al., 2015; McNallym, 2010). The relationship between managerial practices and R&D in explaining firm performance has recently been studied by Nemlioglu and Mallick (2017) and the authors conclude that they are complementary.

Bloom et al., (2017) devise a model that predicts a positive impact of management on firms' performance, a positive relationship between product market competition and management, and a rise in the level and a fall in the dispersion of management with firm age – all results supported empirically. The authors formalize management either as design or as capital (than can be accumulated and depreciated), in both cases entering into the production function. Furthermore, they solve the problem of the firm and provide simulation results for both types of management (design and capital).

In all these empirical results, education of the employees sometimes enters into the explanatory set for output, performance, and productivity measures, as a control to management. This is crucial as productivity is clearly dependent on the skill intensity of the employees. However, firms demand human capital and this demand would depend on output measures and management. This would be a human capital demand approach that has not yet been taken to data. Additionally, it can be conjectured that the management technology also depends on the human capital employed in the firm, not only due to direct participation of employees in some management decisions in modern companies, but also because firms that demand more skilled labor also demand more skilled managers.

We take this alternative avenue to highlight the effect that management has in the skill intensity (or demand) of firms. The contribution closest to ours is Bender et al. (2018), who use a German firms database and find that better-managed firms recruit and retain workers with higher average human capital. The conceptual point of departure is that the relationship between management and productivity is intermediated by the talent of the CEO. This talent of the CEO concept can be enlarged to the culture of the firm, which is shaped by incentive packages offered to both managers and non-manager workers in the firm. This article also estimates a TFP regression that includes both labor quality and wage premium proxies and concludes that they are important in explaining TFP differences. Also they find that when they are excluded, the management variable obtains a higher coefficient, tending to indicate that in that case, there is omitted variable bias.

In our paper, rather than estimating TFP regressions we estimate human capital (skills) demand regressions. To our knowledge this is the first time this is reported in this literature. Apparently, this only consists of solving the firm's problem in order to the human capital demanded. However, due to the controls in the right-hand-side of the regressions, this yields structurally different results. We also analyze the influence of specific components of management on the demand for skills, which we also consider to be a novelty in the literature that relates management quality to measures of firm behavior or performance.

In Section 2 we present descriptive statistics and some empirical evidence of the relationship between human capital (or skills) employed in firms and management practices followed in the same firms. In Section 3 we devise the model building on Bloom et al. (2017) and obtain the human capital demand equations. and present a simple quantitative exercise. In Section 4 we show the regression results in which we estimate the derived theoretical relationship. We also present regressions including specific components of the management index. Finally, in Section 5 we conclude.

2. Descriptive Statistics and Empirical Motivation

In this Section we present descriptive statistics (Table 1) on the main variables used in the paper. As sources we use the WMS data provided by Bloom and Van Reenen (2010), Bloom and Van Reenen (2007], and Bloom et al., (2012a). Figure 1 presents the distribution of both the log (% Employees with a degree) and log (Management).

Table 1: Descriptive Statistics

	Mean	SD	Min	Max
Data from Bloom and Van Reenen (2010)				
log (% employees with a degree)	0.066	0.449	0	4.554
log (Management)	1.061	0.242	0	1.609
log (Capital/employee)	1.406	1.859	-4.48	9.239
Data from Bloom et al. (2012a)				
log (% employees with a degree)	1.655	1.348	-3.912	4.605
log (Management)	1.085	0.219	0.054	1.587
log (Capital/employee)	3.617	1.176	-2.555	9.225
Data from Bloom and Van Reenen (2007)				
log (% employees with a degree)	2.754	0.855	0.598	4.554
log (Management)	1.145	0.265	0.054	1.609
log (Capital/employee)	3.382	0.802	0.261	6.025
log (Wages)	3.633	0.332	2.996	4.605

Figure 1: Distribution of Management Score and Demand for Skills

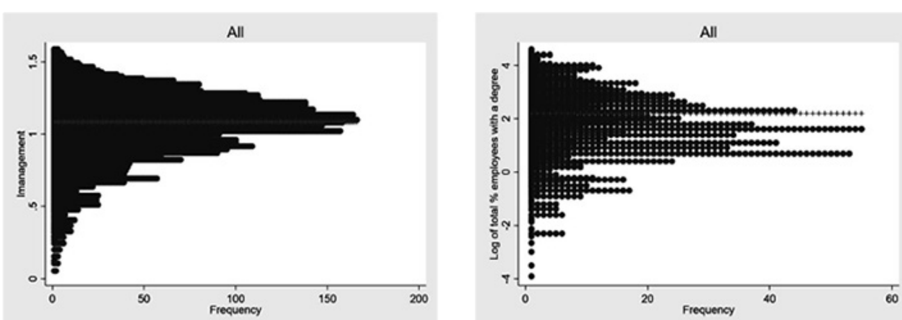
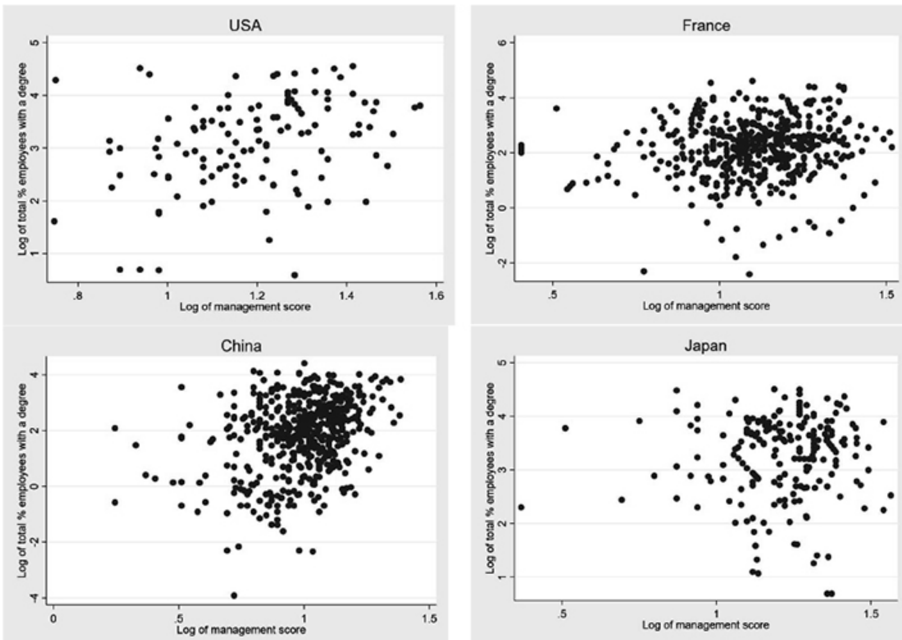


Figure 2 presents scatterplots of the two variables for specific countries. Simple correlations between both variables oscillate significantly from a lower positive correlation of 3% in Japan, to values around 15% for Germany, France and the UK and attains values nearly 31% for China and the USA.

Figure 2: Examples of Scatterplots between Management and Demand for Skills for a Set of Countries



In the Appendix we show regressions using similar specifications to the authors, but with the dependent variable being the *% Employees with a degree*. Despite using relatively similar methods and data, small changes in specifications and data lead to quite different coefficients for management in regressions for human capital (or skills intensity). This calls for the need for some theoretical guidance on the specification of the equation for skills to be estimated. The model in Section 3 provides such guidance.

3. **The Model**

The model builds on Bloom et al. (2017) but is modified to include human capital (or skills) and efficiency wages.

3.1. Setup

The final good technology in each firm is

$$Y_i = F(A_i, H_i, K_i, M_i), \tag{1}$$

where A is technology or Total Factor Productivity (TFP), H is human capital, K is physical capital, and M is Management. The *Management as Technology* perspective assumes that some types of *best practices* of management (e.g. not promoting incompetent employees to senior positions, or collecting some information before making decisions, Taylor's Scientific Management; Lean Manufacturing; Deming's Total Quality Management, incentive pay etc.) increases efficiency. It is obvious that some of these practices are directly linked with the intensity of skills employed and so we can expect that management practices increase the intensity of skills. On the contrary, the *Management as Design* perspective assumes that differences in practices are simply styles optimized to a firm's environment. This means that some practices could increase (or decrease) efficiency depending on this environment. A particular example is purely tenured-based which can lead to a reduction of influence activities but otherwise (or in other firms) reduce output.

Without loss of generality we assume that output Y_i is a real quantity and thus following the *Management as Technology* perspective we use a Cobb-Douglas technology as in [Bloom et al., 2017], extended to allow for human capital and individual effort determining efficiency:[1]

$$Y_i = \left(A_i H_i\right)^a K_i^b M_i^c, \tag{2}$$

with $0 < a, b, c < 1$, $A_i = A_0 e(w_i, w_a)$ denoting that productivity is determined by efficiency in work. This means that the efficiency of work (or effort e) is determined by industry-specific labor market conditions χ, which can be further specified including unemployment rates, u wages in firms that compete for the same skills, w_a and the own wage w. We specify effort as:

$$\begin{cases} \left(\dfrac{w_i - \chi}{\chi}\right)^{\beta} if w > w_a \\ 0 \ otherwise \end{cases} \tag{3}$$

where β measures the concavity of the effort function. While human capital is accumulated outside the firm (by households), physical capital and management are accumulated by the firm, such as:

$$K_{it} = (1 - \delta_k)K_{it-1} + I_{k,it}, \tag{4}$$

$$M_{it} = (1 - \delta_m)M_{it-1} + I_{m,it}, \tag{5}$$

where δ_k and δ_m are depreciation rates of physical capital and management and $I_{k,it}$ and $I_{m,it}$ are investment in both types of capital, with the additional restriction that management capital cannot be sold and so $I_{m,it} \geq 0$. The firms' static problem can be written as follows:

$$Max\ \pi_{H_i, w_i} A_0{}^a \left(\left(\frac{w_i - \chi}{\chi}\right)^{\beta} H_i\right)^a K_i^b M_i^c - w_i H_i \tag{6}$$

[1] As in that paper, we also assume that since firms in our data are typically small in relation to their input and output markets, for tractability we ignore any general equilibrium effects, taking all input prices (for capital, labor, and management) as constant.

Finally, the firm demand for skills or human capital and wage come from the firms' maximization problem (6) in order to human capital and wage:

$$H_i = \frac{aY_i}{w_i} = \frac{a\left(A_i H_i\right)^a K_i^b M_i^c}{w_i} = \frac{aA_0^{\frac{1}{1-a}}\left(\frac{w_i-\chi}{\chi}\right)^{\frac{a\beta}{1-a}} K_i^{\frac{b}{1-a}} M_i^{\frac{c}{1-a}}}{w_i} \tag{7}$$

$$w_i = \frac{\chi}{1-\beta} \tag{8}$$

where equation (7) comes from the equality of the wage and marginal productivity of skills – and the last equality from solving for H_i – and equation (8) comes from the so-called Solow Condition. This yields the following equation for the demand of skills:

$$H_i = \frac{aY_i}{w_i} = \frac{aA_0 H_i^a K_i^b M_i^c}{w_i} = \frac{aA_0^{\frac{1}{1-a}}\left(\frac{w_i-\chi}{\chi}\right)^{\frac{a\beta}{1-a}} K_i^{\frac{b}{1-a}} M_i^{\frac{c}{1-a}}}{w_i} = \frac{aA_0^{\frac{1}{1-a}} K_i^{\frac{b}{1-a}} M_i^{\frac{c}{1-a}}}{\frac{\chi}{1-\beta}\left(\frac{\beta}{1-\beta}\right)^{\frac{a\beta}{1-a}}} \tag{9}$$

3.2. Calibration and a quantitative exercise

We want to infer some quantitative properties of the model and compare them with the econometric estimations we perform in the next section. To that end, we calibrated the model. For the parameters of the production function we assume constant returns to scale in equation (2), setting $a = 0.4$, $b = 0.1$ and $c = 0.5$.[2] The parameters of the efficiency wage setting are $\chi = 1$ and $\beta = 0.5$, assuming a concave function in (3). Depreciation for physical capital and management (as a technology) are in line with the literature (5% for physical capital and 1% for management, assuming that management practices – or culture – depreciates less than physical capital). For the initial levels of physical capital, human capital, and management we use values from the data averages in Table 1. The initial value for output is calculated using equation (2) and assuming $A_0 = 1$. Finally, investment in physical capital assumes a flexible accelerator approach for which we need a real interest rate (assumed to be $r = 0.1$), and the value for accelerator, assumed to be 0.2.[3] In the baseline the investment in Management will be zero, and so, $m_{acc} = 0$. Most of the assumptions will be relaxed in some of the exercises.

[2] These values are in line with the estimated coefficients e.g. in Table 3. Note that small changes in these values, namely the assumption of decreasing returns to scale, do not change the nature of our quantitative results.

[3] In this case investment is given by $I_t = acc_k(\Delta Y/(r + \delta_k))$.

Table 2: Calibration

Calibrated values						
a	B	c	χ	β	δ_k	δ_m
0.4	0.1	0.5	1	0.5	0.05	0.01
Initial Values and Additional variables						
K_0	H_0	M_0	Y_0	r	k_{acc}	m_{acc}
4.08	1.068	2.889	2.009	0.1	0.2	0

In the first exercise (Figure 2(a)) the main force in place is the depreciation rate for physical capital, which makes the series decrease following a higher initial value. In Figure 2(b) we observe the resulting evolution of the series after a one-off positive shock in Management (we introduce a nearly 1/3 increase of the initial value). Output, physical capital, and demand for human capital initially respond positively to the shock but decrease thereafter. Most interesting scenarios happen when we allow for a permanent shock in management allowing for a 20% increase in the score (of the previous period) per period (Figure 2(c)). In this exponential growth case, output and the demand for skills also grow exponentially. After 30 periods the demand for skills rises almost 100 times, at an average period growth rate of 4.6%. Finally in Figure 2(d), we assume a more modest permanent increase in management – 10% increase in the score (of the previous period) per period. Note that in any case the increase in management is always a force in opposition to that of the depreciation effects since there is no exogenous shock in management other than technology.[4] In this last case, this becomes especially visible since the evolution of physical capital is U-shaped. Only after a certain period does the positive effect of management offset and eventually surpass the negative effect of depreciations. This is also visible in the demand for skills, which is much flatter than before. At the end of the 30th period the demand for skills is almost at the same level as the average value of the data, the departing point.

[4] The evolution of investment in physical capital is endogenous.

Figure 3: Simulated Series for Capital, Human Capital, Output, and Management

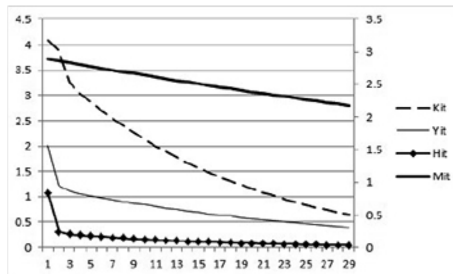


Figure 3(a)

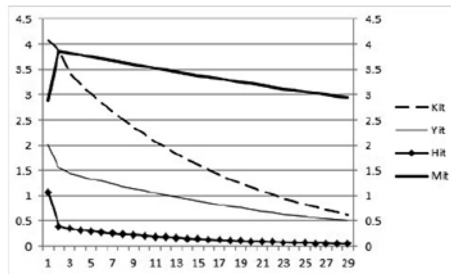


Figure 3(b)

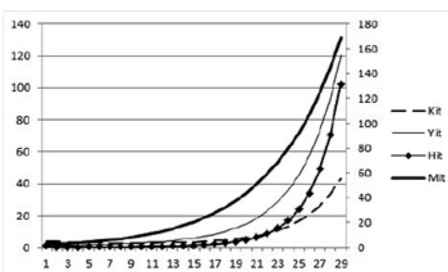


Figure 3(c)

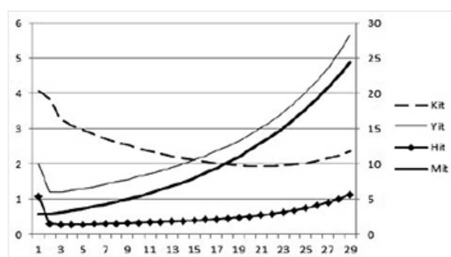


Figure 3(d)

Note: Right-hand scale is for Management.

The effect of management in the demand for skills may be calculated as $\Delta H/\Delta M$. We do that for the first 30 periods. This yields an average value per period of 36.1% in the first (baseline) scenario, 34.7% in the second scenario, 37.9% in the third, and 3.11% in the last one.

4. **Estimation**

We now estimate equation (2) in log form, using the percent of college degree to proxy H_i, capital per employee to proxy K_i, the management index M and general and noise controls as in regressions of Section 2. Specifically, industry dummies proxy the possible effect of industry labor market conditions.

Table 3 shows high significance for coefficients on Management using a log-log specification uncovered by a simple model with efficiency wages, in spite of very different quantitative effects depending on the database used. A 1% increase of Management increases the percentage of college degrees employed from 1.9% to 129.5%. This means that if a firm has 20% of college degree holders in its workforce, a 1% increase in the quality of management index would imply that it will have nearly 24% to nearly 46% level in human capital due to the management quality rise.

These values are consistent with the almost 40% increase in the demand for skills for a 1% increase in management obtained in the simulation of the model we presented above. This leads us to believe that the simple model we devised to highlight the relationship between Management and the demand for skills is particularly useful in predicting realistic quantitative effects. We also learn that differences in estimates may derive from different investment patterns in management (both investment and depreciation rates) that may be present in different databases.

Table 3: Regressions for skills

Dependent variable: log (% employees with a college degree)				
	(1)	(2)	(3)	(4)
log (Management)	0.019***	0.627***	1.295***	0.929***
	(0.006)	(0.097)	(0.279)	(0.000)
log (Capital/employee)	0.001	0.062***	0.016	-0.000***
	(0.001)	(0.022)	(0.036)	(0.000)
log (Wages)	--	--	--	0.000***
				(0.000)
Firms	5085	2927	523	313
Observations	27481	7094	4293	2218

Notes: ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.1 levels, respectively. Standard Standard-errors presented in parentheses are clustered by firm when there are several observations by firm and heteroscedasticity-robust otherwise. Constants and all controls are included in regressions but not shown in the table. Column (1) presents the results of a regression using data from Bloom and Van Reenen (2010). Column (2) presents the results of a regression using data from Bloom et al. (2012a). Column (3) presents the results of a regression using data from Bloom and Van Reenen (2007). Column (4) presents the results of a regression using data from Bloom and Van Reenen (2007), in which we also control for firms' own wages (which are not available in other databases).

4.1. The influence of sub-items of management

The management score is divided into four main dimensions: lean, performance, target, and talent. The first is focused on production processes, the second focuses on how performance is measured and tackled. The third focuses on how the firm defines and interconnects goals between the short and the long run and between financial and nonfinancial goals. Finally, talent captures how the firm implements policies that reward, promote, and attract talents. Those four dimensions may have different effects in the demand for skills. Table 5 shows results in which each of these four dimensions are introduced. Looking at the results we can evaluate the quantitative effects of those four dimensions in the demand for skills. Interestingly, all sub-items help to increase the demand for skills. The most important quantitatively are the target and talent dimensions followed by performance and lean, respectively. It is interesting that a 1% increase in target leads to a 43% to 103% increase in the percentage of college degrees employed, a 1% increase in talent to an increase of nearly 70%, and a 1% increase in performance to an increase of between 30% and 70% increase in the percentage of college degrees employed. Finally, a 1% increase in lean would lead to, at best, a 20% increase in the percentage of college degrees employed.

Another issue that is interesting to be explored is the effect of each of those components maintaining the overall management score as constant. This could indicate to firm which dimension it might wish to act in so as to increase the employment of skills, and also to policy makers that are interested in increasing the skill intensity of the firms. Our results show that in that case increasing target and talent while decreasing performance for a given level of management will increase the demand for skills.[5]

Table 4: Regressions for skills, sub-items

Dependent variable: log (% employees with a college degree)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log (Lean)	0.184***	0.240	--	--	--	--	--	--
	(0.054)	(0.190)						
log (Performance)	--	--	0.296***	0.693***	--	--	--	--
			(0.074)	(0.176)				
log (Talent)	--	--	--	--	0.681***	0.748***	--	--
					(0.090)	(0.273)		
log (Target)	--	--	--	--	--	--	0.429***	1.035***
							(0.077)	(0.183)
Firms	2924	523	2027	523	2927	523	2927	523
Observations	7088	4293	7090	4293	7094	4293	7094	4293

Notes: ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.1 levels, respectively. Standard-errors presented in parentheses are clustered by firm when there are several observations by firm and heteroscedasticity-robust otherwise. Constants and all controls (including log (Capital/employee)) are included in regressions but not shown in the table. Odd Columns present the results of a regression using data from Bloom et al. (2012a). Even Columns present the results of a regression using data from Bloom and Van Reenen (2007). The first database used for regressions in Table 4 does not have information for the sub-items.

[5] Results are available upon request. This means that we obtain significant and positive coefficients for target and talent in regressions in which the (total) management score also enters as covariate, and negative and significant coefficients for performance are obtained in those regressions. Lean becomes nonsignificant in regressions in which the (total) management score also enters as covariate.

5. Conclusion

Research on the influence of management in firms' performance has been focused on productivity measures. Alternatively, our focus is on the influence of management in the demand for skills. We devise a simple firms model highlighting that investment in management as a technology as well as its depreciation may be at the center of the explanation of such a linkage.

Empirical estimations show high significance for coefficients on Management using a log-log specification. A 1% increase of Management increases the percentage of college degrees employed from 1.9% to 129.5%. This means that if a firm has 20% of college degree holders in its workforce, a 1% increase in the quality of management index would imply that it will have nearly 24% to nearly 46%. These values are consistent with the almost 40% increase in the demand for skills for a 1% increase in management obtained in the simulation of the model we presented above. We also present evidence of the influence of the sub-items of Management on skills' demand and discovered that, besides the talent component of Management, target and performance components greatly influence the demand for skills.

References

Bender, S.; Bloom, N.; Card, D.; Van Reenen, J.; Wolter, S. (2018) Management practices, workforce selection, and productivity. *Journal of Labor Economics*, 36(S1).

Bloom, N.; Genakos, C.; Sadun, R.; Van Reenen, J. (2012a) Management practices across firms and countries. *NBER Working Paper Series*, Working Paper 17850.

Bloom, N.; Lemos, R.; Sadun, R.; Scur, D.; Van Reenen, J. (2014) The new empirical economics of management. *NBER Working Paper Series*, Working Paper 20102.

Bloom, N.; Lemos, R.; Sadun, R.; Scur, D.; Van Reenen, J. (2016) International data on measuring management practices. *American Economic Review*, 106(5), 152-56.

Bloom, N.; Lemos, R.; Sadun, R.; Van Reenen, J. (2015) Does management matter in schools? *Economic Journal*, 125(584), 647-674.

Bloom, N.; Sadun, R.; Van Reenen, J. (2012b) Americans do it better: US multinationals and the productivity miracle. *American Economic Review*, 102(1), 167-201.

Bloom, N.; Sadun, R.; Van Reenen, J. (2017). Management as a technology? Harvard Business School Strategy Unit Working Paper No. 16-133 and Stanford University Graduate School of Business Research Paper No. 16-27.

Bloom, N.; Schweiger, H.; Van Reenen, J. (2012c) The land that lean manufacturing forgot? *Economics of Transition*, 20(4), 593-635.

Bloom, N.; Van Reenen, J. (2007) Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics*, 122(4).

Bloom, N.; Van Reenen, J. (2010) Why do management practices differ across firms and countries. *Journal of Economic Perspectives*, 24(1), 203-224.

Gibbons, R.; Henderson, R. (2013) *The Handbook of Organizational Economics*. Princeton: Princeton University Press.

McNally, G. (2010) Scholarly productivity, impact, and quality among academic psychologists at group of eight universities. *Australian Journal of Psychology*, 62(4), 204-215.

Meagher, K.; Strachan, R. (2013) Gaussian process: A smooth and flexible approach to estimating index complementarities in organizational economics. *Organizational Economics Proceedings*, 2(1).

Milgrom, P.; Roberts, J. (1990) The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review*, 80(3), 511-28.

Nemlioglu, I.; Mallick, S. (2017) Do managerial practices matter in innovation and firm performance relations? New evidence from the UK. *European Financial Management*, 23(5), 1016-1061.

Appendix

Estimations in Selected Databases

In this appendix, we present regressions based on panel data from Bloom and Van Reenen (2010) and Bloom et al. (2012a).

Table A.1: Regressions for skills with data from Bloom and Van Reenen (2010)

Dependent variable: % Employees with a college degree					
	(1)	(2)	(3)	(4)	(5)
Management	0.099***	0.073***	0.021***	0.024***	0.036***
	(0.014)	(0.016)	(0.007)	(0.008)	(0.008)

Notes: ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.1 levels, respectively. Standard-errors presented in parentheses are clustered by firm when there are several observations by firm and heteroscedasticity-robust otherwise. Data from Bloom and Van Reenen (2010). Why Do Management Practices Differ across Firms and Countries? *Journal of Economic Perspectives*, Vol. 24, No. 1. First column includes log (Sales/Employee) as covariate, 4399 firms and 13611 observations. Second column includes log (Sales/Employee), country & industry dummies, 3657 firms and 10392 observations. Column (3) adds general controls and noise controls and log (Capital/Employee), 3391 firms and 9696 observations. Column (4) drops log (Sales/Employee) and log (Capital/Employee) but includes Profitability (ROCE), and the three types of controls, 2491 firms and 8650 observations. Column (5) includes all previous controls simultaneously and 1542 firms, and 5283 observations. General controls include firm-level controls for log(average hours worked) and log(firm age) and noise controls include 78 interviewer dummies, the seniority and tenure of the manager who responded, the day of the week the interview was conducted, the time of day the interview was conducted, the duration of the interviews, and an indicator of the reliability of the information as coded by the interviewer.

Table A.2: Regressions for skills with data from Bloom et al. (2012a)

Dependent variable: % Employees with a college degree									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Management	0.292***	0.270***	0.191***	0.197***	0.322***	0.270***	0.212***	0.234***	0.234***
	(0.024)	(0.025)	(0.015)	(0.015)	(0.036)	(0.035)	(0.036)	(0.035)	(0.035)

Notes: ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.1 levels, respectively. Standard-errors presented in parentheses are clustered by firm when there are several observations by firm and heteroscedasticity-robust otherwise. Data from Bloom et al. (2012a). *Academy of Management Perspectives*, Vol. 26, No. 1. Columns (1) and (2) are for non-managers and use 5407 observations. Columns (3) and (4) are for managers and use 7559 observations. Column (5) includes log (Sales/Employee) as covariate, 2927 firms and 7094 observations. Column (6) includes log (Sales/Employee), country and industry dummies, 2927 firms and 7094 observations. Column (7) adds general controls -- without firm age -- and noise controls and Log(Capital/Employee), 2901 firms and 7000 observations. Column (8) drops log (Sales/Employee) and log (Capital/Employee) but includes Profitability (ROCE) and the three types of controls, using 2,901 firms and 7,000 observations. Column (9) includes the three types of controls and sales growth using 2,901 firms, and 7,000 observations.

Tomé Lima
Service Delivery Manager at FundsDLT, Luxembourg
tometavaresdelima@gmail.com
orcid.org/ 0000-0001-5846-9301
Helder Sebastião
Univ Coimbra, CeBER, Faculty of Economics
helderse@fe.uc.pt
orcid.org/ 0000-0002-5687-3818
DOI: https://doi.org/10.14195/2183-203X_57_3

Native Market Factors for Pricing Cryptocurrencies

Fatores Intrínsecos de Mercado para Avaliação das Criptomoedas

Tomé Lima
Helder Sebastião

Received for publication: October 10, 2023
Revision accepted for publication: December 15, 2023

ABSTRACT
The cryptocurrency market has been growing frantically in number of cryptocurrencies, online exchanges, and market capitalization, which has amplified the need for comprehensive and robust pricing models. Using a database of all eligible cryptocurrencies listed on the CoinMarketCap website, we study the relationship between returns and several potential pricing factors, such as size (market capitalization), momentum, liquidity, and maturity. The analysis was conducted from December 27, 2013, to December 29, 2020, using weekly data for 3'667 cryptocurrencies. Results point out that portfolios of cryptocurrencies with smaller market capitalization, higher reversal, lower liquidity, and lower maturity tend to offer higher returns. The 5-factor model that additionally includes illiquidity and maturity performs better than the 3-factor model previously proposed in the literature, meaning that illiquidity and maturity significantly help capture the cross-sectional cryptocurrency risk premia. The 5-factor model presented seems robust to different procedures to construct portfolios and factors.
Keywords: Bitcoin; cryptocurrencies; asset pricing; factor models.

JEL Classification: G12; G14; G15.

Aknowledgments: This work has been funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020 and Project RiskBigData PTDC/MAT-APL/1286/2021.

1. INTRODUCTION

The growth of the cryptocurrency market in terms of number of cryptocurrencies, online exchanges, and market capitalization has attracted more effective and potential individual and institutional investors. Consequently, the demand for financial studies has also increased, resulting in an exponential growth in the empirical finance literature applied to cryptocurrencies. Until 2017, the attention was focused on a few major cryptocurrencies, such as Bitcoin, Ethereum, Litecoin, Tether and Ripple. More recent papers consider bigger samples formed by more cryptocurrencies and longer periods.

For traditional financial markets, namely the stock market, several studies have attempted to identify the main pricing factors. The Capital Asset Pricing Model (CAPM) that considers just one factor – the market portfolio – is the most simple and well-known of such models. On this topic, Fama (1970), Fama and French (1993), Carhart (1997), Fama and French (2012), and Fama and French (2015) are pivotal references in the related literature. In the cryptocurrency market, this analysis is only beginning with an additional difficulty as some of the factors designed for the stock market are not applicable. Shen et al. (2020) construct a 3-factor model for cryptocurrencies, which encompasses market, size, and momentum factors. Because the book-to-market factor does not apply to cryptocurrencies, the size factor has been constructed using size and momentum. More accurately, this last factor should be called reversal, as it seems that bad (good) past returns tend to be followed by good (bad) returns in the cryptocurrency market. Shahzad et al. (2020) elaborate on this model, adding a contagion factor.

This paper addresses the issue of what market intrinsic factors are priced in the cryptocurrencies market. The main objective of this research is twofold. First, analyze several market features that may drive the prices of cryptocurrencies. Second, use this information to derive a factor pricing model.

The principal data and methodological novelties that this study brings to the literature are the following:

a) Handling a comprehensive dataset of cryptocurrencies, employing all the information in the CoinMarketCap website from April 30, 2013, to December 29, 2020.
b) Consideration of several features of the cryptocurrencies' ecosystem, namely market return, size, momentum, and, most importantly, liquidity and maturity.
c) Application of four different methodologies to construct the portfolios, namely, sequential and intersecting double-sort equally and value-weighted portfolios.
d) Presentation of a 5-factor model that outperforms both the CAPM and the 3-factor model of Shen et al. (2020).

The remainder of this paper is organized as follows. Section 2 presents the arguments supporting additional factors in the pricing model and develop the additionally hypotheses contextualized in the literature. Section 3 explains the raw dataset, filtering procedures, and data aggregation. Section 4 presents the formulas used to compute the financial features of cryptocurrencies, and the methodology to construct the factors and portfolios used in the regressions' framework. Section 5 shows the main results and Section 6 performs some robustness checks. Section 7 concludes the paper.

2. Literature Review and Hypothesis Development

An important strand of financial literature on cryptocurrencies focuses on the weak form of market efficiency, according to which the price system should contain all the relevant information on historical prices and other market-related variables, so that future prices cannot be predicted using past information.

Several studies, such as Urquhart (2016), Nadarajah and Chu (2017), Bariviera (2017), Bariviera et al. (2017), mainly conclude that Bitcoin was weakly inefficient, although it tends to be more efficient as the market evolves and matures.

More recently, other studies began testing the efficiency of other cryptocurrencies besides Bitcoin. Wei (2018) analyses 456 cryptocurrencies in 2017, when the value of the cryptocurrency market was skyrocketing. The author uses the Amihud illiquidity ratio (Amihud, 2002) to sort the cryptocurrencies into five groups and then applies the tests used in Urquhart (2016). Wei (2018) argues that, as more active and informed traders enter the market, liquidity increases while volatility decreases, creating fewer arbitrage opportunities, and hence, highly liquid cryptocurrencies tend to be more efficient. In the same line of thought, Brauneis and Mestel (2018) use 73 cryptocurrencies from August 31, 2015, to November 30, 2017, and conclude that as the liquidity of cryptocurrencies increases, they became less predictable and therefore more efficient. Al-Yahyaee et al. (2020) analyze six cryptocurrencies with the highest market capitalization during the period August 7, 2015, to July 3, 2018, showing that informational efficiency is directly linked to liquidity and that efficiency tends to increase as the market matures.

Several studies have tried to directly identify variables that have a significant relationship with the returns of cryptocurrencies, among these variables stand out size, momentum, trading volume, volatility, and maturity (Liu et al., 2022). Kyriazis & Prassa (2019) analyse 846 cryptocurrencies from April 1, 2018, to January 31, 2019, when the market capitalization of cryptocurrencies was decreasing. They argue that during downward market movements, cryptocurrencies with higher market capitalization are also the ones with higher liquidity. The reasoning is that during bearish periods, investors in most markets tend to prefer assets with higher market capitalization and lower volatility. Brauneis et al. (2020) conclude that liquidity of cryptocurrencies is mostly independent from other financial markets and depends mainly on intrinsic volatility and trading volume. Balcilar et al. (2017) show that trading volume can be used to predict Bitcoin returns but only when the market is performing around the median. Burggraf and Rudolf (2020), using data on 1'000 cryptocurrencies from April 28, 2013, to November 1, 2019, show that higher volatility produces higher returns.

In a nutshell, these studies indicate that volatility is higher in more illiquid and younger cryptocurrencies. As risk should be rewarded by the market, then we formulate the following hypotheses:

H1: Illiquidity increases the returns of cryptocurrencies; hence the illiquidity factor may be measured by a portfolio formed by a long position in illiquid cryptocurrencies and a short position in liquid cryptocurrencies.

H2: Maturity decreases the returns of cryptocurrencies; hence the maturity factor may be measured by a portfolio formed by a long position in younger cryptocurrencies and a short position in older cryptocurrencies.

3. Data and Preliminary Analysis

The dataset was retrieved from https://coinmarketcap.com, which is one of the most complete and reliable sources of information on cryptocurrencies. The legitimacy of this website derives from its use by many financial studies on cryptocurrencies. This website uses objective criteria according to which cryptocurrencies and online exchanges must comply to be listed.[1]

The sample covers the period from April 30, 2013, to December 29, 2020. The raw dataset is formed by 5'763 cryptocurrencies. For each cryptocurrency we retrieved the daily close prices, trading volume, and market capitalization, in USD, recorded at 00:00:00 UTC. According to CoinMarketCap, the close prices are volume-weighted index prices and daily volumes are the simple sum of the trading volume considering several listed online exchanges.

Given that we use the complete set of listed cryptocurrencies, it is important to mention that the data does not suffer from survival bias, as some cryptocurrencies did not reach the last day in the sample. The number of cryptocurrencies increased steadily from 7 on April 30, 2013, to 4'073 on December 29, 2020, but during the overall period covered, 5'763 cryptocurrencies were listed, hence 1'690 cryptocurrencies ceased to exist or were removed from the CoinMarketCap listing. This means that only around 70% survived until December 29, 2020.

The second step in preparing the dataset was filtering the raw data. This was conducted using three filter rules: (1) Trading volume is missing from April 30, 2013, to December 27, 2013. So, the sampling period begins in this last date. The period between these dates is only used to compute the maturity of cryptocurrencies. (2) Some cryptocurrencies had missing days, probably due to communication failures between the exchanges and the CoinMarketCap website. If a particular day was missing, the gap was fulfilled by linear interpolation. We proceeded in this way when there was a maximum of three days missing in a row. Larger gaps, mainly due to provisionally listing on the CoinMarketCap website, were treated as if the cryptocurrency was nonexistent during that period. (3) When a cryptocurrency was added to CoinMarketCap, usually the information on market capitalization for the first few days is not complete or has clear mistakes. These days were ignored for these cryptocurrencies until they had information on all variables of interest.

After applying these filters, we end up with 3'667 cryptocurrencies, 2'562 days, corresponding to 366 weeks. This daily database was then aggregated weekly, using Wednesday-to-Wednesday prices, volumes, and market capitalizations.[2]

---

[1] The complete listings criteria can be accessed at https://support.coinmarketcap.com/hc/en-us/articles/360043659351-Listings-Criteria.

[2] Besides data on the cryptocurrency market, we also collected data on the risk-free rate. Following the literature, and since cryptocurrencies data are expressed in USD, we collected from https://fred.stlouisfed.org/data the yield-to-maturity of 1-month US Treasury bills.

4. Methodology

This section explains the construction of the time series of returns and other features, namely size, illiquidity, momentum, and maturity, for each cryptocurrency. It explains the construction of portfolios and presents some preliminary results that point out how to construct the pricing factors. Finally, it presents the procedures used to compute the pricing factors and the factor models.

4.1. Returns and other features

Since cryptocurrencies are studied cross-sectionally in aggregated terms, i.e., using portfolios, we use discrete returns which are aggregable in the asset space. The close-to-close prices were used to compute the weekly returns of cryptocurrency *i* as:

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-7}}{P_{i,t-7}}, \tag{1}$$

where $P_{i,t}$ and $P_{i,t-7}$ represent the close price on Wednesday t and seven days before, respectively.

The series of returns present massive extreme values, with some cryptocurrencies having returns over 10^4. To *winsorize* the outliers but still maintain the main features of the data, namely volatility, we used an interquartile distance to identify and rescale outliers. We consider as an outlier any observation outside the interval of $[p_{25} - k(p_{75} - p_{25}), p_{75} + k(p_{75} - p_{25})]$, where p_{25} and p_{75} are the 25th and 75th percentiles, respectively, and k is a multiplier factor. We tested several multipliers, $k = 1.5$, 3, 4.5, 6 and 7, and decided to use $k = 6$. Using this criterium, 99.81% and 89.96% of cryptocurrencies have less than 5% and 1% of outliers, respectively, which were rescaled to the limits of the above interval.

Size was simply proxied by the market capitalization.

For the momentum we followed Shahzad et al. (2020) and Shen et al. (2020), which conclude that the best strategy, i.e., the one with the higher t-statistic, results from forming buy-sell portfolios based on the previous returns for a one-time holding period. This means constructing the portfolios at time $t - 7$, based on the returns of the cryptocurrencies from $t - 14$ to $t - 7$, and holding it until Wednesday t, which translates into

$$Mom_{i,t} = R_{i,t-7}, \tag{2}$$

where, $R_{i,t-7}$ is the weekly return of cryptocurrency i at $t - 7$.

Brauneis et al. (2021) explore high and low frequency data for Bitcoin and Ethereum, testing different liquidity measures, and concluding that one of the best measures to describe the liquidity of cryptocurrencies was the Amihud illiquidity ratio (Amihud, 2002). Hence, illiquidity was measured by this ratio, which assesses the price impact of 1USD of trading volume on the returns. Theoretically, the ratio ranges from 0 (most liquid) to $+\infty$ (most illiquid). For a given cryptocurrency i, the illiquidity ratio was computed as:

$$Ill_{i,t} = \frac{1}{7}\sum_{\tau=t-7}^{t}\frac{\left|R_{i,\tau}\right|}{V_{i,\tau}}, \tag{3}$$

where $R_{i,\tau}$ and $V_{i,\tau}$ are the arithmetic return and the volume traded in USD at day τ, respectively.

For measuring the maturity of a cryptocurrency, we considered the number of weeks with valid data from its launching until day t. To compute this measure, we use all the data available since April 30, 2013. On this date only seven cryptocurrencies were listed, hence for all other cryptocurrencies, there is no measurement error.

4.2. Portfolios

We consider four features: size (market capitalization), momentum, measured by the previous weekly return, illiquidity, measured by the Amihud illiquidity ratio, and maturity, measured by the number of weeks since launching. These portfolios are constructed on $t - 7$ and held until t. Table 1 enables a first glance at the importance of each feature and the way that portfolios should be combined to compute the pricing factors.

Table 1: Weekly mean returns of quintile portfolios

	Quintiles				
	1	2	3	4	5
Size	0.0914	0,0506	0.0308	0.0167	0.0130
Momentum	0.0294	0.0064	0.0055	0.0182	0.0106
Illiquidity	0.0128	0,0018	0.0066	0.0327	0.0915
Maturity	0.0132	0.0150	0.0030	0.0191	-0.0073

Notes: This table presents the weekly mean returns of value-weighted quintile portfolios. Each week, all cryptocurrencies were sorted by a given feature (size, measured by market capitalization, momentum, measured by the previous weekly return, illiquidity, measured by the Amihud illiquidity ratio, and maturity, measured by the number of weeks since launching) and are partitioned into quintiles. Then, the value-weighted portfolio, where the weight of each cryptocurrency is given by its relative market capitalization, is computed for each quintile. The sample is from January 1, 2014, to December 29, 2020 (365 weeks).
Source: Authors' own calculations.

The patterns in Table 1 suggest that portfolio returns increase inversely with size, momentum, liquidity, and maturity. The size and momentum effects are in accordance with the literature (see, for instance, Shahzad et al., 2020; Shen et al., 2020; Liu et al.,2022). The reported illiquidity and maturity effects support our hypotheses H1 and H2, respectively.

To form double-sorted portfolios of cryptocurrencies we use a sequential procedure. This procedure is as follows: (1) At each $t - 7$, all cryptocurrencies are sorted based on the market capitalization (i.e., size) and are grouped into quintiles, (2) within each size quintile, cryptocurrencies are then sorted by the second feature and once again clustered into quintiles, (3) we then form value-weighted portfolios, using market capitalization as the weighting scheme, and compute their returns from $t - 7$ to t, which are then used to compute the excess returns in relation to the risk-free rate (1-month US Treasury bill). Hence, according to each pair size/other feature we obtain 25 value-weighted portfolios. This approach is different from Fama and French (1993, 2012, 2015), that form 25 value-weighted portfolios by intersecting quintiles from a sort on size with the quintiles from an independent sort on the second feature. Our procedure produces portfolios with the same number of cryptocurrencies (except the last quintile portfolios which include the remaining cryptocurrencies, if the total number is not a multiple of 5), whilst the Fama-French approach gives portfolios with a variable number of cryptocurrencies. Another approach, such as the one used by Carhart (1997), is to construct equally weighted portfolios.

The weekly excess returns of these portfolios are presented in Table 2. Most portfolios excess returns are significant at the 1% level, and portfolios with cryptocurrencies of small, illiquid, with lower momentum (higher reversal) and lower maturity have higher excess returns. From all the different portfolios, it is quite visible that portfolios with smaller size offer higher excess returns.

Table 2: Average excess returns of sequential double sorted value-weighted portfolios

	Size and momentum					
	Down - 1	2	3	4	Up - 5	D - U
Small - 1	0.3120***	0.0926***	0.0590***	0.0540***	-0.0324***	0.3443***
2	0.1882***	0.0552***	0.0365***	0.0301***	-0.0466***	0.2346***
3	0.1187***	0.0313***	0.0243***	0.0222***	-0.0354***	0.1540***
5	0.0738***	0.0131*	0.0089	0.0077	-0.0180**	0.0916***
Big -5	0.0041	0.0033	0.0111*	0.0176**	0.0151	-0.0111
S - B	0.3078***	0.0891***	0.0477***	0.0362***	-0.0476***	
	Size and illiquidity					
	Liquid - 1	2	3	4	Illiquid - 5	I - L
Small - 1	0.0476***	0.0877***	0.0888***	0.1149***	0.1529***	0.1051***
2	0.0210**	0.0464***	0.0401***	0.0573***	0.1007***	0.0795***
3	0.0151*	0.0233***	0.0217***	0.0277***	0.0733***	0.0580***
4	0.0108	0.0121*	0.0155**	0.0172**	0.0303***	0.0193**
Big - 5	0.0132**	-0.0019	0.0136	0.0034	0.0152	0.0019
S - B	0.0343***	0.0895***	0.0750***	0.1113***	0.1375***	

	Size and maturity					
	Young - 1	2	3	4	Old - 5	Y – O
Small - 1	0.0850***	0.0895***	0.0975***	0.0767***	0.1052***	-0.0203
2	0.0519***	0.0565***	0.0453***	0.0442***	0.0523***	-0.0006
3	0.0277***	0.0418***	0.0330***	0.0274***	0.0224**	0.0051
4	0.0181**	0.0238***	0.0116	0.0158**	0.0145	0.0034
Big - 5	0.0129**	0.0098	0.0097	0.0125	-0.0022	0.0150*
S - B	0.0719***	0.0796***	0.0876***	0.0640***	0.1072***	

Notes: In each week $t-7$, all active cryptocurrencies were sorted into quintiles by size (market capitalization) and then, within these quintes were sorted by a second feature. The excess returns of week t were computed using the yield-to-maturity of the 1-month US Treasury bills. Portfolios are updated on a weekly basis (there are 365 weekly observations, from January 1, 2014, to December 29, 2020). The last column is obtained by subtracting in each week the portfolios in quintiles 1 and 5. Line S-B is obtained in each column by subtracting the line Big from line Small. ***, **, * indicates significance at the 1%, 5 % and 10% level, respectively.
Source: Authors' own calculations.

4.3. Pricing factors and models

The pricing factors are built on the previous portfolios, conditional on the pair size/other feature. For the market factor, like in CAPM, we consider the value-weighted total market index (*MKT*) using all the cryptocurrencies in our filtered database as:

$$MKT_t = \sum_{i=1}^{N} R_{it} \frac{MarketCap_{it}}{\sum_{i=1}^{N} MarketCap_{it}}, \tag{4}$$

where R_{it} is the return and $MarketCap_{it}$ is the market capitalization of cryptocurrency i at the beginning of week t , and N is the number of cryptocurrencies.

Since cryptocurrencies do not have a book-value, to construct the size factor, we follow the approach suggested by Shen et al. (2020) and use momentum as the second sort. From these two sorts, and similar to Fama and French (2015), we divide the size sort by percentile [0%, 10%] (Small) and percentile [90%, 100%] (Big), and the momentum sort by percentile [0%, 30%] (low momentum, denoted by Down), percentile]30%, 70%[(Medium momentum) and percentile [70%, 100%] (higher momentum, denoted by Up). Then we intersect the size and momentum partitions, creating six value-weighted portfolios, respectively, SD, SM, SU, BS, BM, and BU.

From the evidence presented in Table 1 and Table 2, Small portfolios offer higher returns than Higher portfolios, hence the size factor is defined as Small minus Big (*SMB*):

$$SMB_t = \frac{SD_t + SM_t + SU_t}{3} - \frac{BD_t + BM_t + BU_t}{3}. \tag{5}$$

For the remaining factors, we proceeded in the same way but dropping the medium interval on the second feature. Our factors, were, respectively, Down momentum minus Up momentum (*DMU*), Illiquid minus Liquid (*IML*), and Young minus Old (*YMO*). That is:

$$DMU_t = \frac{BD_t + SD_t}{2} - \frac{BU_t + SU_t}{2}, \tag{6}$$

$$IML_t = \frac{BI_t + SI_t}{2} - \frac{BL_t + SL_t}{2}, \tag{7}$$

$$YMO_t = \frac{BY_t + SY_t}{2} - \frac{BO_t + SO_t}{2}, \tag{8}$$

With all the portfolios and factors constructed, we proceeded with the estimation of the factor models using Ordinary Least Square (OLS).

The first model only considers the market factor, similar to CAPM, with the market portfolio proxied by the value-weighted market index, *MKT*.

$$R_{p,t} - Rf_t = a + b_1 (MKT_t - Rf_t) + \varepsilon_{p,t}, \tag{9}$$

where $R_{p,t}$, Rf_t, and MKT_t are the return of portfolio p, the risk-free interest rate, and the market return at time t, respectively.

As in Shen et al. (2020), the 3-factor model is defined by:

$$R_{p,t} - Rf_t = a + b_1 (MKT_t - Rf_t) + b_2 SMB_t + b_3 DMU_t + \varepsilon_{p,t}, \tag{10}$$

where *SMB* and *DMU* are respectively the size and momentum factors previously defined. Our more encompassing model is a 5-factors model, defined as:

$$R_{p,t} - Rf_t = a + b_1 (MKT_t - Rf_t) + b_2 SMB_t + b_3 DMU_t + b_4 IML_t + b_5 YMO_t + \varepsilon_{p,t}, \tag{11}$$

where *IML* and *YMO* are the illiquidity and maturity factors, respectively.

As in Fama and French (2012), we defined the Sharpe ratio as:

$$SR = \left(a'\Omega^{-1}a\right)^{\frac{1}{2}} \tag{12}$$

where a is the column vector of the intercepts of the regressions and Ω is the covariance matrix of the error terms.

5. Main Empirical Results

Table 5 presents a summary of the average statistics for the CAPM, 3-factor, and 5-factor models. This table highlights that the 5-factor model improves on the CAPM and on the

3-factor model. The average absolute intercept decreases and the GRS statistic (Gibbons et al., 1989) on the null hypothesis that the intercepts are jointly equal to zero, although still significant at the 1% level, decrease substantially. The average standard error of the intercepts decreases and the adjusted R^2 increases. Notice that although the additional factors are important in explaining the returns of cryptocurrencies, the market factor is undoubtedly the most important one.

Table 5: Summary statistics on CAPM, 3-factor and 5-factor models

	|a|	R^2	s(a)	SR	GRS
CAPM	0.0319	0.3432	0.0076	1.0521	15.124***
3-factor	0.0214	0.4074	0.0093	0.9181	7.0763***
5-factor	0.0204	0.4170	0.0094	0.7663	5.0777***

Notes: This table presents the summary statistics from regressions on CAPM, 3-factor and 5-factor models. Each column corresponds to the average statistics for the regressions on sequential double-sort value-weighted portfolios. |a| is the average absolute intercept, R^2 is the average adjusted determination coefficient, s(a) is the average standard error of the intercepts. SR is the Sharpe ratio computed according to Equation (12). GRS is the statistics on the null hypothesis that the intercepts are jointly zero (Gibbons et al., 1989). The significance at the 1%, 5% and 10% is denoted by ***, **, *, respectively. Regressions were performed using 365 weekly observations, from January 1, 2014, to December 29, 2020.
Source: Authors' own calculations.

6. **Robustness Checks**

The results presented in the previous section may be sensitive to the way that factors and portfolios are constructed, hence we conduct several robustness checks on the CAPM, 3-factor and 5-factor models.

Procedure 1 – The same sequential double-sort procedure but instead of using value-weighted portfolios when grouping the cryptocurrencies, we consider equally-weight portfolios.

Procedure 2 – For each pair size/another feature, portfolios are created using Fama and French (1993, 2012, 2015) procedure, that is, by intersecting the independent sort on size with an independent sort on another feature. From these intersections we formed both (2.1) value-weighted and (2.2) equally weighted portfolios. Table 6 shows the summary statistics of Procedure 1 and Procedure 2.

Table 6: Robustness checks on the portfolio construction

	Procedure 1 – Sequential double-sort equally weighted portfolios				
	|a|	R^2	s(a)	SR	GRS
CAPM	0.0371	0.3484	0.0074	1.1521	17.966***
3-factor	0.0227	0.4268	0.0089	0.7575	5.4514***
5-factor	0.0221	0.4286	0.0091	0.7502	5.2012***

	Procedure 2.1 – Double-sort intersection value-weighted portfolios				
	|a|	R^2	s(a)	SR	GRS
CAPM	0.0313	0.3249	0.0081	1.0408	14.833***
3-factor	0.0215	0.3856	0.0100	0.8865	6.6357***
5-factor	0.0206	0.3960	0.0101	0.7600	5.0160***
	Procedure 2.2 – Double-sort intersection equally weighted portfolios				
	|a|	R^2	s(a)	SR	GRS
CAPM	0.0355	0.3266	0.0078	1.1855	18.977***
3-factor	0.0219	0.4006	0.0094	0.9079	7.7592***
5-factor	0.0214	0.4039	0.0096	0.9149	7.6476***

Notes: This table presents the summary statistics for regressions on CAPM, 3-factor and 5-factor models considering different ways to construct the portfolios. Alternatives are the sequential double-sort but with equally weighted portfolios, the double-sort intersection value-weighted portfolios of Fama and French (1993, 2012, 2015), and the double-sort intersection but with equally weighted portfolios. Each column corresponds to the average statistics for the regressions. |a| is the average absolute intercept. R^2 is the average adjusted determination coefficient, s(a) is the average standard error of the intercepts. SR is the Sharpe ratio computed according to Equation (12). GRS is the statistics on the null hypothesis that all the intercepts for a set of regressions are jointly zero (Gibbons et al., 1989). The significance at the 1%, 5% and 10% is denoted by ***, **, *, respectively. Regressions were performed using 365 weekly observations, from January 1, 2014, to December 29, 2020.
Source: Authors' own calculations.

Procedure 3 –On the previous factors we used the percentile [0%, 10%] as small size cryptocurrencies and the interval [90%, 100%] as big size cryptocurrencies. Here we use percentiles [0%, 50%] and]50%, 100%], i.e., the median to divide the cryptocurrencies into Small and Big. The breakpoints on the second feature are the same as before using the intervals [0%, 30%],]30%, 70%[and [70%, 100%]. Using these factors, we estimate the 3 models for the following portfolios: (3.1) sequential double-sort value-weighted, (3.2) sequential double-sort equally weighted, (3.3) double-sort intersection value-weighted, and (3.4) double-sort intersection equally weighted. Table 7 shows the summary statistics of Procedure 3.

Table 7: Robustness checks on the portfolio and factor constructions

	Procedure 3.1 – Sequential double-sort value-weighted portfolios				
	|a|	R^2	s(a)	SR(a)	GRS
CAPM	0.0319	0.3432	0.0076	1.0521	15.124***
3-factor	0.0271	0.4828	0.0079	1.1510	12.935***
5-factor	0.0247	0.5093	0.0082	0.9796	8.7113***

	Procedure 3.2 - Sequential double-sort equally weighted portfolios				
	|a|	R^2	s(a)	SR(a)	GRS
CAPM	0.0371	0.3484	0.0074	1.1521	17.966***
3-factor	0.0315	0.4921	0.0077	1.0112	10.378***
5-factor	0.0295	0.5079	0.0080	1.0130	9.4491***
	Procedure 3.3 – Double-sort intersection value-weighted portfolios				
	|a|	R^2	s(a)	SR(a)	GRS
CAPM	0.0313	0.3249	0.0081	1.0408	14.833***
3-factor	0.0271	0.4579	0.0086	1.1264	12.366***
5-factor	0.0246	0.4846	0.0088	0.9625	8.4596***
	Procedure 3.4 – Double-sort intersection equally weighted portfolios				
	|a|	R^2	s(a)	SR(a)	GRS
CAPM	0.0355	0.3266	0.0078	1.1855	18.977***
3-factor	0.0304	0.4627	0.0082	1.1039	12.329***
5-factor	0.0287	0.4796	0.0085	1.1220	11.500***

Notes: This table presents the summary statistics from regressions on CAPM, 3-factor and 5-factor models considering different ways to construct the portfolios and to construct the pricing factors. Now, factors are constructed using the median to divide the cryptocurrencies into Small and Big. The breakpoints on the second attribute are kept as before using the intervals [0%, 30%],]30%, 70%[and [70%, 100%]. The alternatives for the portfolios are the sequential double-sort with equally and value-weighted portfolios, the double-sort intersection with equally and value-weighted portfolios. Each column corresponds to the average statistics of the regressions. |a| is the average absolute intercept for a set of regressions, R^2 is the average adjusted determination coefficient, s(a) is the average standard error of the intercepts, and SR is the Sharpe ratio computed according to Equation (12). GRS is the statistics on the null hypothesis that all the intercepts for a set of regressions are jointly zero (Gibbons et al., 1989). The significance at the 1%, 5% and 10% is denoted by ***, **, *, respectively. Regressions were performed using 365 weekly observations, from January 1, 2014, to December 29, 2020.
Source: Author's own calculations.

The results of the several alternative procedures are similar to the ones of the baseline framework, implying that our main results and inferences are robust to the procedures used to construct the portfolios and pricing factors. These results also reinforce the claim that adding liquidity and maturity as pricing factors improves the 3-factor model of Shen et al. (2020) and, in fact, this is especially true when using the median as the partition point for the size factor.

7. **Conclusions**

This study explores several pricing factors of the cryptocurrencies market, for the period from December 27, 2013, to December 29, 2020, using weekly frequency. The methodology is like the one used for the stock market by Fama and French (1993, 2012, 2015), with some nuances on the portfolio and factor constructions. Noticeably, our baseline approach, contrary to Fama and French (2015) and Shen et al. (2020), who produce the value-weighted portfolios by intersecting two independent sorts, is a sequential double-sort procedure that produces portfolios with the same cardinality. However, our main results are not sensitive to the way that portfolios or even pricing factors are constructed.

We were able to identify two additional pricing factors: illiquidity and maturity. Clearly the returns of cryptocurrencies are directly related to the evolution of the overall market, the most important pricing factor. However, there is compelling evidence that cryptocurrencies with lower market capitalization (small size), more illiquid, with higher reversals, and less mature present higher returns.

Our 5-factor pricing model considers the market portfolio, size (Small minus Big – SMB), momentum (Down minus Up – DMU), illiquidity (Illiquid minus Liquid – IML), and maturity (Young minus Old (YMO). The inclusion of illiquidity and maturity improves the results in relation to the 3-factor model of Shen et al. (2020).

We should highlight that we are only dealing with native factors of the cryptocurrency market, i.e., factors that use the information intrinsic to the market. Other external factors such as the investor's attention, proxied for instance by Google searches may be important as it seems to be the case for Bitcoin (see, for instance, Kristoufek, 2015, Dastgir et al., 2019, Anastasiou et al., 2021).

References

Al-Yahyaee, K. H.; Mensi, W.; Ko, H.-U.; Yoon, S.-M.; Kang, S. H. (2020) Why cryptocurrency markets are inefficient: The impact of liquidity and volatility. *The North American Journal of Economics and Finance*, 52, 101168._

Amihud, Y. (2002) Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31-56._

Anastasiou, D.; Ballis, A.; Drakos, K. (2021) Cryptocurrencies' price crash risk and crisis sentiment. *Finance Research Letters*, 101928.

Balcilar, M.; Bouri, E.; Gupta, R.; Roubaud, D. (2017) Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*, 64, 74-81.

Bariviera, A. F. (2017) The inefficiency of Bitcoin revisited: A dynamic approach. *Economics Letters*, 161, 1-4.

Bariviera, A. F.; Basgall, M. J.; Hasperué, W.; Naiouf, M. (2017) Some stylized facts of the Bitcoin market. *Physica A: Statistical Mechanics and Its Applications*, 484, 82-90.

Brauneis, A.; Mestel, R. (2018) Price discovery of cryptocurrencies: Bitcoin and beyond. *Economics Letters*, 165, 58-61.

Brauneis, A.; Mestel, R.; Theissen, E. (2020) What drives the liquidity of cryptocurrencies? A long-term analysis. *Finance Research Letters*, 101537._

Brauneis, A.; Mestel, R.: Riordan, R.; Theissen, E. (2021) How to measure the liquidity of cryptocurrency markets? *Journal of Banking & Finance*, 124, 106041.

Burggraf, T.; Rudolf, M. (2020) Cryptocurrencies and the low volatility anomaly. *Finance Research Letters*, 101683.

Carhart, M. M. (1997) On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57-82.

Dastgir, S.; Demir, E.; Downing, G.; Gozgor, G.; Lau, C.K.M. (2019). The causal relationship between Bitcoin attention and Bitcoin returns: Evidence from the Copula-based Granger causality test'. *Finance Research Letters*, 28, 160-164.

Fama, E. F. (1970) Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383-417._

Fama, E. F.; French, K. R. (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.

Fama, E. F.; French, K. R. (2012) Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105(3), 457-472.

Fama, E. F.; French, K. R. (2015) A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

Gibbons, M. R.; Ross, S. A.; Shanken, J. (1989) A test of the efficiency of a given portfolio. *Econometrica*, 57(5), 1121-1152.

Kyriazis, N. A.; Prassa, P. (2019) Which cryptocurrencies are mostly traded in distressed times? *Journal of Risk and Financial Management*, 12(3), 135.

Kristoufek, L. (2015) What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PLoS one*, 10(4) e0123923.

Liu, Y.; Tsyvinski, A.; Wu, X. (2022) Common risk factors in cryptocurrency. *Journal of Finance*, 77(2), 1133-1177.

Nadarajah, S.; Chu, J. (2017) On the inefficiency of Bitcoin. *Economics Letters*, 150, 6-9.

Shahzad, S. J. H.; Bouri, E.; Ahmad, T.; Naeem, M. A.; Vo, X. V. (2020) The pricing of bad contagion in cryptocurrencies: A four-factor pricing model. *Finance Research Letters*, 101797.

Shen, D.; Urquhart, A.; Wang, P. (2020) A three-factor pricing model for cryptocurrencies. *Finance Research Letters*, 34, 101248.

Urquhart, A. (2016) The inefficiency of Bitcoin. *Economics Letters*, 148, 80-82.

Wei, W. C. (2018) Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, 168, 21-24.

Pedro Bação
Univ Coimbra, CeBER, Faculty of Economics
pmab@fe.uc.pt
orcid.org/0000-0002-3340-1068
José Diogo Silva
Univ Coimbra, Faculty of Economics
d.agostzz11@gmail.com
DOI: https://doi.org/10.14195/2183-203X_57_4

Inflation in Portugal Through the Lens of the Fair Model

A Inflação em Portugal da Perspetiva do Modelo de Fair

Pedro Bação
José Diogo Silva

Received for publication: October 1, 2023
Revision accepted for publication: December 12, 2023

ABSTRACT
This study delves into the dynamics of inflation in Portugal, employing a cost-push model as the analytical framework. The model is estimated using data from 2000Q2 to 2020Q1, a period predating the onset of the COVID-19 pandemic and the surge in inflation. We then produce forecasts spanning from 2020Q2 to 2023Q2. The forecasts hint strongly at a structural break during this latter period, implying that the model offers insufficient representation of inflation dynamics in Portugal. We conclude with a discussion of the model's strengths and limitations in understanding inflation dynamics, shedding light on critical aspects that impact its explanatory power.
Keywords: Cost push; econometric modeling; Fair model; forecasting; inflation.

JEL Classification: E17; E31; E37.

Aknowledgment: This work has been funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020.

1. Introduction

In 2022-2023, inflation was one of the highest worries, if not the highest, around the world (e.g., World Economic Forum, 2022) and also in Portugal (e.g., Villalobos, 2023). From 2010 to 2021, average inflation in Portugal was 1%; in 2022 inflation was 7.8%. Economic policymakers have been urged to take action against it. However, the remedies prescribed have been met with controversy. For example, the Italian, Portuguese, and Spanish governments publicly criticized the ECB's decision to increase interest rates in June 2023 (Albert, 2023).

The question of what drives inflation is not only important for designing economic policy but also a highly contentious issue. A standard model of inflation is the cost-push model. In the cost-push model, inflation is viewed as determined by the evolution of production costs. Ray Fair's macroeconometric model of the USA (Fair, 2018) is a prominent example of this approach to modelling inflation. The cost-push model may be a good modelling choice in the current context, given that the recent inflation has been associated with rising production costs, in particular energy costs, coupled with the turmoil caused by the pandemic and the war (Lane, 2022). In this paper, we use the price level equation from Fair's model (equation 10 in that model) as the starting point for analyzing the evolution of inflation in Portugal in recent years. The results presented in Silva (2023) indicate that this model may perform better than alternative models in the Portuguese case.

In Section 2, we present Fair's price level equation and describe the procedure we used for constructing the corresponding time series for Portugal. In Section 3, we present the empirical results. We first estimate the model using data for 2000Q2-2020Q1, i.e., the period before the Covid-19 pandemic and the rise in inflation. We use the model to produce forecasts for 2020Q2-2023Q2. The forecasts, along with the results of a Chow test, strongly suggest that a structural break occurred during the latter period. Section 4 discusses the results and concludes the paper.

2. The Model and the Data

2.1. Fair's price level equation

The price level equation in Fair's model is the following:

$$\log PF_t = \beta_1 + \beta_2 \log PF_{t-1} + \beta_3 a_t + \beta_4 \log PIM_t + \beta_5 UR_t + \beta_6 t + \beta_7 CB_t + \beta_8 TB_t + \varepsilon_t \quad (1)$$

PF is the price deflator for nonfarm firm sales, *a* is a measure of wage costs (discussed below), *PIM* is the import price deflator, *UR* is the unemployment rate, *t* is time, *CB* is a dummy variable that represents a break in the intercept, and *TB* represents a break in the linear trend.

To use this equation for modelling inflation in Portugal, we must specify the Portuguese time series to replace the US time series. Note that we use seasonally adjusted data to remove seasonality, which would otherwise introduce additional variation. Note also that the model is based on quarterly data; whenever the original data were monthly, we converted them by calculating the mean value. The data sources are shown in Table 1. We use the consumer price index (CPI) for *PF*, the import price deflator (the ratio of nominal to real exports, from the national accounts) for *PIM*, and the unemployment rate for the population aged 16 to 74 years old for *UR*.

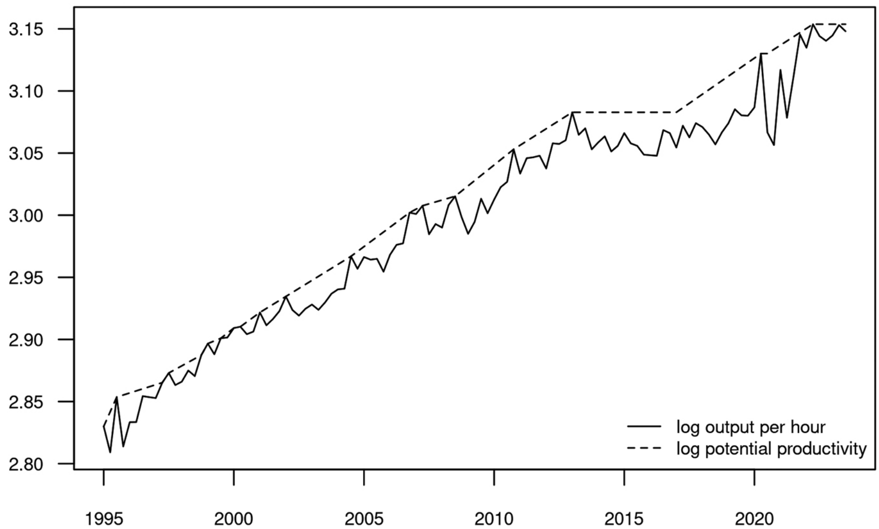
Table 1: Data sources

Time series	Sample
Imports of goods and services (current prices; quarterly)	1995Q1-2023Q3
Imports of goods and services (chain linked volume data; quarterly)	1995Q1-2023Q3
Gross domestic product at market prices (chain linked volume data; quarterly)	1995Q1-2023Q3
Total hours worked (quarterly)	1995Q1-2023Q3
Compensation of employees (current prices; quarterly) (S.1 Total economy)	1999Q1-2023Q2
Consumer price index (CPI, Base-2012)	1948M1-2023M11
Unemployment rate (Seasonally adjusted) in percentage of active population aged between 16 and 74 years old	1998M2-2023M10
Value Added Tax	1986M1-2023M11

Source: Statistics Portugal (INE) and economias.pt (in the case of VAT).

The measure of wage costs is the difference between the logarithm of the wage rate of the firm sector and the logarithm of potential productivity. We approximate the wage rate with the ratio between the compensation of employees and the number of hours worked, both from the national accounts. Fair adjusts the wage rate to incorporate the employer social security tax rate, but the Portuguese equivalent has been constant over the sample period. As for potential productivity, Fair determines it by choosing a set of peak dates and interpolating the logarithm of output per hour worked between each pair of consecutive peak dates. We proceed in a similar manner. First, we compute output per hour worked as the ratio of real GDP to the number of hours worked, both from the national accounts. Second, we determine the sequence of maxima in output per hour worked; this provides a list of candidate peak dates. Third, we eliminate candidate peak dates that result in short peak-to-peak intervals (we set the minimum size to six quarters). We then compute an interpolated series in each interval; the interpolation minimizes a function that weights the variation in the interpolated series and the deviation from the actual output per hour worked. Finally, we reintroduce some of the previously excluded candidate peak dates and recompute the interpolated series. Namely, we reintroduce those which are necessary to ensure that potential productivity is never below actual output per hour worked. The result is shown in Figure 1.

Figure 1: Potential productivity



2.2. Breaks

Fair's model of inflation also includes a deterministic component with breaks. The behaviour of inflation in Portugal has likewise been characterized by apparent breaks. Given the sample size, we allowed for five breaks and used the following algorithm to determine the break dates:

1. Start with an empty set of break dates, *A*.
2. For each date *t* in the sample:

    2.1. Add *t* to the set *A* to form a new candidate set *B*.
    2.2. Create the "break variables" corresponding to set *B*.
    2.3. Estimate a model with the break variables (besides the intercept and the linear trend) for the log of CPI (the "price level") and for the first difference of log CPI ("inflation").
    2.4. Save the sum of squared residuals (sum the squared residuals from both models).

3. Update set *A* to include the break date with the lowest sum of squared residuals in step 2.
4. If the number of break dates in *A* is less than five, return to step 2, otherwise stop.

Besides the variables used by Fair and the break variables, we included in the model the value added tax rate. One of the measures taken by the Portuguese Government to deal with the increase in inflation was to reduce the VAT rate on certain (essential) items. In the empirical model we use the normal VAT rate (currently 23%).

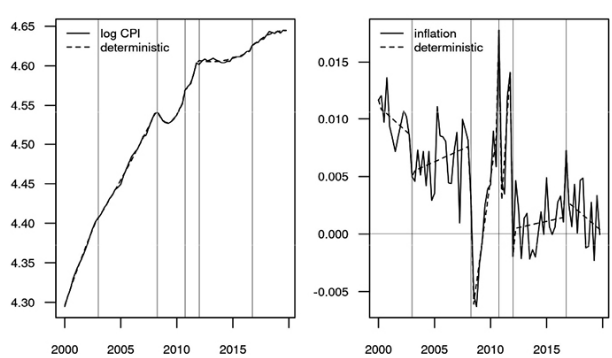
The "break variables" for each break date t_i were the following:

1) $\text{dbreak}_t = 1$ if $t = t_i$; $\text{dbreak}_t = 0$ otherwise
2) $\text{cbreak}_t = 1$ if $t \geq t_i$; $\text{cbreak}_t = 0$ otherwise
3) $\text{tbreak}_t = t - t_i + 1$ if $t \geq t_i$; $\text{tbreak}_t = 0$ otherwise
4) $\text{qtbreak}_t = (t - t_i + 1)^2$ if $t \geq t_i$ and $< t_{i+1}$; $\text{qtbreak}_t = 0$ otherwise

The quadratic terms were restricted to each interval, while the changes in the intercept and in the linear trend affected all the following observations. We made this choice because quadratic trends do not appear to be as common as linear trends in the data. Furthermore, the model for inflation used the dbreak_t, cbreak_t and tbreak_t variables, while the model for the price level used the cbreak_t, tbreak_t and qtbreak_t variables. In other words, we allow for a quadratic trend in the level, but not in the first difference. Recall that a quadratic trend in the level suggests a linear trend in the first difference, a linear trend in the level suggests a change in the intercept in the first difference, and a change in the intercept in the level suggests a dummy variable for the corresponding break date in the first difference.

In the determination of break dates, as well as in the initial analysis presented below, we used the sample 2000Q2-2020Q1. This excludes the most recent period (where inflation took off). The reason is that we want to see how a model estimated on the initial part of the sample behaves in the final part of the sample. We also exclude the initial part of the sample because there appeared to be another break in that period. Consequently, the procedure leads to the following break dates: 2003Q2, 2008Q3, 2011Q1, 2012Q2, and 2017Q1. The fitted values from the deterministic models corresponding to these break dates for the price level and for inflation are plotted in Figure 2. The vertical lines identify the break dates.

Figure 2: Fitted values of the deterministic models with breaks



3. Empirical Analysis

3.1. The pre-pandemic period

We started by estimating the following version of Fair's price level equation:

$$\log CPI_t = \beta_1 + \beta_2 \log CPI_{t-1} + \beta_3 a_t + \beta_4 \log PIM_t + \beta_5 UR_t + \beta_6 VAT_t + (\textit{break variables}) + \varepsilon_t \qquad (2)$$

However, the Breusch-Godfrey autocorrelation test rejected the null hypothesis of no autocorrelation up to order four. Therefore, we added one lag of the explanatory variables (apart from the break variables) to the model. We rewrote the model in an equivalent way in which the dependent variable is inflation and only levels lagged once and first differences appear as explanatory variables:

$$\Delta \log CPI_t = \beta_1 + \beta_2 \log CPI_{t-1} + \beta_3 a_{t-1} + \beta_4 \log PIM_{t-1} + \beta_5 UR_{t-1} + \beta_6 VAT_{t-1} + \beta_7 \Delta \log CPI_{t-1} + \beta_8 \Delta a_t + \beta_9 \Delta \log PIM_t + \beta_{10} \Delta UR_t + \beta_{11} \Delta VAT_t + (\textit{break variables}) + \varepsilon_t \qquad (3)$$

This version of the model is closer to the idea that inflation is essentially the result of an adjustment of the price level towards an "equilibrium" level defined in some way. In our model, a natural assumption is that the equilibrium level depends on the lagged levels (except the lagged price level) and on the break variables (except the observation-specific dummies). This amounts to assuming that inflation can be written as

$$\Delta \log CPI_t = \beta_2 (\log CPI_{t-1} - \textit{equilibrium level}_{t-1}) + (\text{short} - \text{run elements}) \qquad (4)$$

where β_2 should be between -1 and 0, for the adjustment towards the equilibrium level to actually occur. The "short-run elements" are assumed to have a zero mean, except in the case of observation-specific dummies.

The inclusion of the break variables greatly increases the number of explanatory variables. We use a stepwise regression procedure to eliminate statistically insignificant variables. The stepwise regression procedure removes the explanatory variable with the highest p-value exceeding 10%, re-estimates the model without that variable, and continues this process until no variable remains with a p-value exceeding 10%. The resulting model is in Table 2.

Table 2: Estimated model, 2000Q2-2020Q1

	coefficient	std. error	t-ratio	p-value
Constant	0.4009	0.0356	11.27	0.0000***
log CPI_{t-1}	-0.1270	0.0121	-10.47	0.0000***
log PIM_{t-1}	0.0296	0.0110	2.696	0.0091***
Δ log PIM_t	0.0642	0.0145	4.416	0.0000***
VAT_{t-1}	0.1204	0.0464	2.596	0.0119**
ΔVAT_t	0.2828	0.0336	8.419	0.0000***
qtbreak_20002	0.0001	0.0000	3.597	0.0007***
qtbreak_20032	0.0000	0.0000	6.189	0.0000***
cbreak_20032	0.0073	0.0023	3.233	0.0020***
tbreak_20083	0.0012	0.0002	7.762	0.0000***
dbreak_20083	0.0069	0.0013	5.189	0.0000***
qtbreak_20111	0.0006	0.0000	19.01	0.0000***
tbreak_20111	-0.0013	0.0002	-6.847	0.0000***
cbreak_20111	-0.0063	0.0013	-4.978	0.0000***
dbreak_20111	0.0094	0.0010	9.54	0.0000***
qtbreak_20122	0.0000	0.0000	2.706	0.0088***
cbreak_20122	0.0054	0.0011	5.101	0.0000***
dbreak_20122	-0.0037	0.0008	-4.518	0.0000***
cbreak_20171	0.0089	0.0029	3.035	0.0036***
dbreak_20171	0.0023	0.0007	3.544	0.0008***

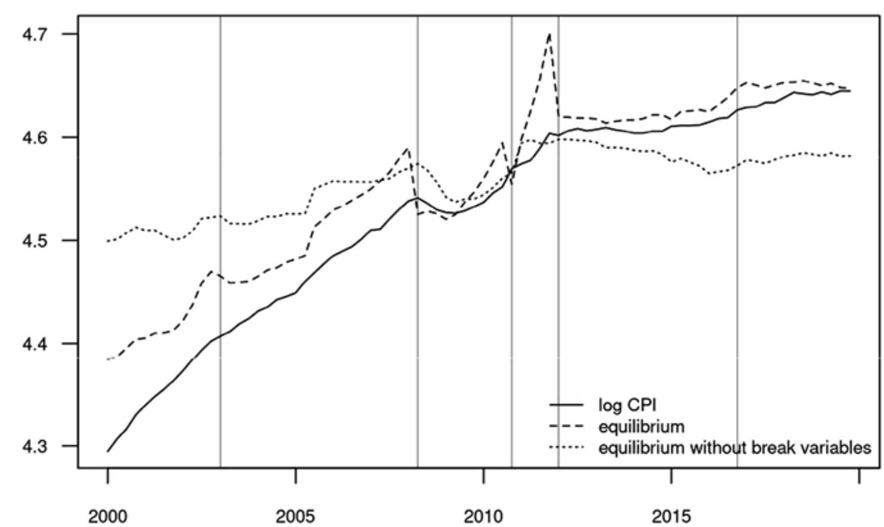
Mean dependent var	0.004525	S.D. dependent var	0.004742
Sum squared resid	0.00021	S.E. of regression	0.001872
R-squared	0.8816	Adjusted R-squared	0.844107
Log-likelihood	400.4406	Akaike criterion	-760.8812
Schwarz criterion	-713.2407	Hannan-Quinn	-741.7808
rho	-0.231809	Durbin-Watson	2.454941

Note: The dependent variable is inflation (in equation 3).

The signs of the coefficients are as expected. Namely, lagged CPI has a negative coefficient, implying a speed of adjustment towards the equilibrium level of 13% per quarter. The price of imported goods has a positive coefficient, reflecting the importance of imported inflation. The VAT rate also has a positive coefficient, suggesting that it influences the equilibrium level and therefore inflation during the adjustment period.

However, the model in Table 2 is missing two of the variables in Fair's equation: the wage pressure and the unemployment rate. This would imply that the Portuguese price level is driven by imported inflation, VAT adjustments, and a trend with breaks. In our view, this is likely an indication that the Fair model, in the format estimated here, is not well-suited to the Portuguese case. If we compare the evolution of the estimated equilibrium price level with and without the break variables – see Figure 3 – the conclusion is that the break variables account for most of the increase in the equilibrium price level. In fact, the change, from 2000Q2 to 2020Q1, in the equilibrium price level computed without the break variables is roughly one third of the change in the equilibrium price level computed with the break variables. The fact that the model does not tell us what makes the break variables behave as they do means that the model leaves a significant part of the evolution of the price level essentially unexplained.

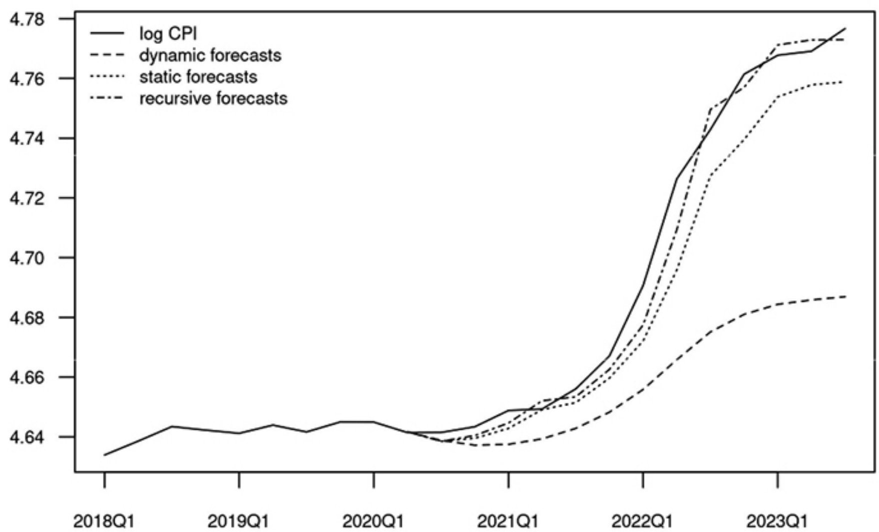
Figure 3: Equilibrium price level



3.2. Forecasts using the pre-pandemic model

Milton Friedman wrote that "The true test of a scientific theory-of a set of propositions about a class of observable phenomena-is whether it works, whether it correctly predicts the consequences of changes in conditions." (Friedman, 1968, p.15). Although in a different context, let us then give our model a chance to prove its worth in forecasting the subsequent (2020Q2-2023Q3) evolution of the price level in Portugal. We compute three sets of forecasts: dynamic, static and recursive. All of them employ the actual evolution of the explanatory variables during the forecast period. Dynamic forecasts use the forecast for the price level in the next period to forecast the price level in the period after. Static forecasts use the actual next-period price level to forecast the price level in the period after. Recursive forecasts also use the actual next-period price level to forecast the price level in the period after; in addition, the model is re-estimated in each period. The forecasts are plotted in Figure 4.

Figure 4: Forecasts 2020Q2-2023Q3



The dynamic forecasts signal an uptick in inflation (driven by import prices), but clearly understate its magnitude. Static forecasts perform better, but do not fully eliminate the gap to the actual price level. Recursive forecasts apparently track the evolution of the price level reasonably well. However, the fact that they perform much better than the other forecasts suggests that the model's parameters have changed, i.e., that another break has occurred. A Chow test, with the break date set to the beginning of the forecast period, 2020Q2, corroborates that view (test statistic $F(5,69) = 21.7086$ with p-value 0.0000).

Application of the procedure for selecting the break dates described in subsection 2.2. points to 2021Q4 as the new break date. This is the quarter before Russia invaded Ukraine. It is also a period when restrictive measures were once again imposed as a reaction to an increase in Covid-19 cases.

4. Conclusion

The analysis of the empirical model presented in this paper, based on Fair's (2018) cost-push model, sheds light on both its strengths and limitations in capturing the dynamics of inflation in Portugal. While the model incorporates explanatory variables such as import prices and VAT adjustments, and performs well in fitness measures, it falls short in fully explaining the observed fluctuations in inflation. There are several dimensions to this failure.

Firstly, the model omits variables such as wage cost pressure and the unemployment rate, which, within the framework of the model, are expected to be important for understanding inflation dynamics. Secondly, the dynamic, static, and recursive forecasts for the most recent years suggest the presence of a structural break. This hypothesis is supported by a standard Chow test. Our analysis suggests that the new break occurred in 2021Q4, coinciding with significant geopolitical and pandemic-related events. Thirdly, the dominant driver of the price level in the model is a deterministic component with breaks, but the determinants of this component remain unexplained. This leaves a notable gap in our understanding of what drives inflation.

References

Albert, E. (2023) Signs of a political revolt are brewing around the European Central Bank. Le Monde (July 6). Retrieved from https://www.lemonde.fr/en/opinion/article/2023/07/06/signs-of-a-political-revolt-are-brewing-around-the-european-central-bank_6044074_23.html

Fair, R. C. (2018) *Macroeconometric modeling: 2018*. Retrieved from https://fairmodel.econ.yale.edu/mmm2new1.htm

Friedman, M. (1968) *Dollars and Deficits*. Prentice-Hall.

Lane, P. R. (2022) Inflation Diagnostics. Retrieved from https://www.ecb.europa.eu/press/blog/date/2022/html/ecb.blog221125~d34babdf3e.en.html

Silva, J. D. A. (2023) Uma Análise Econométrica da Inflação em Portugal. Work Project, (Master in Economics, Faculty of Economics of the University of Coimbra). Retrieved from https://estudogeral.uc.pt/handle/10316/109607

Valor do IVA em Portugal ao longo dos anos (2018) Retrieved from https://www.economias.pt/valor-do-iva-em-portugal-ao -longo-dos-anos

Villalobos, L. (2023) Inflação e custo de vida identificados como os principais problemas do país. *Público* (February 26).

World Economic Forum (2022) Inflation tops the list of the world's biggest worries but there's plenty more on our minds, survey suggests (October 21). Retrieved from https://www.weforum.org/agenda/2022/10/worries-inflation-economy-concerns/

1 2 9 0

IMPRENSA DA
UNIVERSIDADE
DE COIMBRA
COIMBRA UNIVERSITY PRESS