

**SOURCES OF ECONOMIC POLICY  
UNCERTAINTY IN THE EURO AREA:  
A READY-TO-USE DATABASE**

2023

**BANCO DE ESPAÑA**  
Eurosistema

Documentos Ocasionales  
N.º 2315

Andrés Azqueta-Gavaldón, Marina Diakonova,  
Corinna Ghirelli and Javier J. Pérez

**SOURCES OF ECONOMIC POLICY UNCERTAINTY IN THE EURO AREA: A READY-TO-USE  
DATABASE**

## SOURCES OF ECONOMIC POLICY UNCERTAINTY IN THE EURO AREA: A READY-TO-USE DATABASE (\*)

Andrés Azqueta-Gavaldón

ARCTURIS DATA

Marina Diakonova

BANCO DE ESPAÑA

Corinna Ghirelli

BANCO DE ESPAÑA

Javier J. Pérez

BANCO DE ESPAÑA

(\*) We want to thank Javier Vallés Liberal for his comments. This paper should not be reported as representing the views of the Banco de España, nor Arcturis Data. The views expressed are those of the authors and do not necessarily reflect those of the Banco de España, the Eurosystem or Arcturis Data. The indicators can be downloaded from this webpage: [link](#). E-mail addresses: andres-azqueta@gmail.com, marina.diakonova@bde.es, corinna.ghirelli@bde.es, javier.perez@bde.es.

Documentos Ocasionales. N.º 2315

August 2023

<https://doi.org/10.53479/33155>

The Occasional Paper Series seeks to disseminate work conducted at the Banco de España, in the performance of its functions, that may be of general interest.

The opinions and analyses in the Occasional Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet on its website at: <http://www.bde.es>.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

© BANCO DE ESPAÑA, Madrid, 2023

ISSN: 1696-2230 (on-line edition)

## Abstract

In this paper, we build a publicly-available database of economic policy uncertainty (EPU) indicators based on the methodology proposed by Azqueta-Gavaldón, Hirschbühl, Onorante and Saiz (2023), which uses topic modelling techniques to identify distinct components of EPU. This database is regularly updated and can be accessed on the Banco de España's website. Currently, the dataset covers the four largest countries in the euro area, namely Spain, Italy, France, and Germany. Our data coverage is continually expanding to include more euro area countries. Additionally, we compute the aggregated EPU indexes for the euro area. This comprehensive dataset and the resulting euro area indexes provide valuable tools for researchers, policymakers and analysts to assess and monitor the dynamics of economic policy uncertainty in real time.

**Keywords:** economic policy uncertainty, euro area, machine learning, Latent Dirichlet Allocation, word embeddings.

**JEL classification:** D80, E20, E66, G18.

## Resumen

Construimos una base de datos de indicadores de incertidumbre de política económica (*economic policy indicators*: EPU, por sus siglas en inglés) de acceso público basada en la metodología propuesta por Azqueta-Gavaldón, Hirschbühl, Onorante y Saiz (2023), que utiliza técnicas de *topic modelling* para identificar los distintos componentes de incertidumbre. Esta base de datos se actualiza periódicamente y es accesible a través de la página web del Banco de España. Actualmente, los indicadores abarcan los cuatro países más grandes de la zona euro: España, Italia, Francia y Alemania. Además, agregando los indicadores nacionales de estos cuatro países, calculamos un indicador EPU para la zona euro. Estamos en el proceso de ampliar la cobertura de datos para construir indicadores EPU para más países de la UEM. Este conjunto de datos y los índices derivados para la zona euro proporcionan valiosas herramientas a investigadores, responsables políticos y analistas para evaluar y supervisar la dinámica de la incertidumbre de la política económica en tiempo real.

**Palabras clave:** incertidumbre acerca de las políticas económicas, zona euro, aprendizaje automático, Asignación Latente de Dirichlet, *word embeddings*.

**Códigos JEL:** D80, E20, E66, G18.

## Contents

Abstract 5

Resumen 6

1 Introduction 8

2 Data and Methods 10

3 Economic policy uncertainty in the euro area 14

4 Conclusions 18

References 19

Appendix 20

## 1 Introduction

Economic Policy Uncertainty (EPU) has gained increasing attention in recent years as a measure of the level of uncertainty surrounding economic policy decisions. The EPU index was first introduced by Baker et al. (2016) and has since been used widely in research to explore the impact of policy uncertainty on economic outcomes. Baker et al. (2016) computed EPU indexes for the most important European countries (Italy, France, Germany and Spain) as well as an European EPU index which is based on the press coverage of these four aforementioned countries. In this spirit and following this procedure, Ghirelli et al. (2019) have developed a new EPU index for Spain that improves the original EPU indicator for Spain by widening the press coverage and refining the keywords.<sup>1</sup>

This construction of the EPU index is not, however, without limitations, particularly when it comes to the specific approach used to create the index (see Naboka-Krell, 2023). To surmount these methodological challenges recent literature has proposed a new set of EPU indicators based on topic modelling techniques: e.g. Azqueta-Gavaldon (2017), Azqueta-Gavaldón et al. (2023), Naboka-Krell (2023).<sup>2</sup> This is the approach taken in this paper. Here we construct the EPU indexes proposed by Azqueta-Gavaldón et al. (2023) based on topic modelling techniques for the four largest euro area countries (Spain, Italy, France, and Germany), as well as for the euro area as a whole.

We improve the original Azqueta-Gavaldón et al. (2023)'s indicators in two main dimensions. First, we enlarge the time coverage backwards, going as far back as 1996 depending on the data availability for each country.<sup>3</sup> Second, we widen the newspaper coverage. For Spain, we are already considering seven sources, in line with the newspapers used by Ghirelli et al. (2019). For the rest of Europe, the coverage is currently similar to Azqueta-Gavaldón et al. (2023), while some beta versions are already available to include seven sources for each country.<sup>4</sup> Our indicators are available in a publicly-available database that is regularly updated at the website of the Bank of Spain: see link. At the moment of writing this paper, our indicators span until 2019, due to technical problems in the download of the raw textual data. We are currently working to solve this issue, and plan to release the updated and complete version of the indicators in the following months.

For the construction of the indexes we closely follow Azqueta-Gavaldón et al. (2023) and apply Latent Dirichlet Allocation (LDA), a topic modelling technique that allows us to capture the underlying topics and themes that contribute to economic policy uncertainty. LDA is an unsupervised machine learning technique that retrieves components of policy

---

<sup>1</sup> All these indicators are regularly updated in the [PolicyUncertainty](#) webpage.

<sup>2</sup> For a more detailed discussion on the advantages of topic modelling, see Azqueta-Gavaldón et al. (2023) and its references therein.

<sup>3</sup> We gather data from 1997 for Italy, Germany, and Spain, and from 1996 for France. By contrast, Azqueta-Gavaldón et al. (2023) consider the period from 2000 onwards.

<sup>4</sup> Andres-Escayola et al. (2022) show that increasing the number of newspaper sources improves the robustness of the resulting EPU indexes.



uncertainty endogenously, without the need to come up with ad hoc dictionary ex-ante. This allows us to obtain a more nuanced and detailed understanding of the sources of economic policy uncertainty. Based on the LDA output, we select eight topics that best describe sources of policy uncertainty in the European context: fiscal; monetary; political; geopolitical; trade/manufacturing; European regulation; domestic regulation; and energy. We then compute eight indexes for each country that represent these sources of uncertainty. In addition, for each country our overall index of economic uncertainty results from aggregating the aforementioned source-specific uncertainty indexes. Finally, we construct the full set of uncertainty indicators (topic-specific uncertainty indicators as well as the overall EPU indicator) for the euro area by averaging the country-specific indicators. In this case the country-specific indicators are weighted by country GDP relative to the euro area GDP.

Consistent with the literature, our euro area EPU indicator displays increases in periods of economic distress such as the Iraq war, the great financial crises, the EU debt crisis, and the Covid crisis. Overall, we find that the EPU average has remained higher for all countries after the Great Financial Crisis, which is consistent with the findings of Baker et al. (2016). For example, the monetary policy component of euro area EPU (EPU-MPU), which is constructed based on the frequency of media references to monetary policy uncertainty in a given country, increases significantly during times of economic distress, such as the beginning of the Great Financial Crisis in 2008-2009, the Brexit Referendum in 2016, and to a lesser extent the Covid-19 crisis. Besides, the geopolitical component of euro area EPU (EPU-GPU), which is constructed based on the frequency of media references to geopolitical uncertainty in a given country, shows significant spikes during times of significant geopolitical unrest, such as the Iraq war in 2003, the Arab Spring in 2011, the Russian invasion of Crimea in 2014, and its recent invasion of Ukraine.

As Azqueta-Gavaldón et al. (2023) show, using LDA to create the EPU indexes can provide a more nuanced and detailed understanding of the sources of economic policy uncertainty. This could have important implications for policymakers and researchers interested in understanding the dynamics of policy uncertainty and its impact on economic outcomes.

The rest of the paper is structured as follows: Section 2 describes the data and the methodology. Section 3 presents the indicators for Italy, Spain, Germany, Spain, and the euro area. In addition, for the case of Spain, we compare our EPU index with the closest ones existing in the literature: the original EPU index by Baker et al. (2016) and the new (improved) EPU index by Ghirelli et al. (2019). Note both of them are dictionary-based, while ours is based on topic modelling techniques.

## 2 Data and Methods

To construct the economic uncertainty indicators, we rely on the Dow Jones news archive. The Banco de España has access to a subset of the European press of Dow Jones data, and receives monthly updates of the raw news data, which allows us to update our indicators on a weekly basis. In particular, the European press data we are considering is limited to the following Dow Jones categories: Commodity/Financial Market News (mcat), Corporate/Industrial News (ccat), Economic News (ecat), Politics/International Relations (gpir), Editorials (nedi), and Commentaries/Opinions (nedc).<sup>5</sup>

### Media Outlet

Following the coverage of Azqueta-Gavaldón et al. (2023) and Ghirelli et al. (2019) for each country, we base our analysis on the following national newspapers:

- Spain: El Mundo (1995), Expansión (1995), Cinco Días (1996), La Vanguardia (1997), ABC (1997), El País (2001), El Economista (2014).
- France: Le Monde (1990), Le Figaro (1997).
- Germany: Süddeutsche Zeitung (1995), Die Welt (2004), Handelsblatt (2013).
- Italy: La Stampa (1996), Corriere della Sera (1997), La Repubblica (2005).

This selection represents between 24% to 69% of the daily circulation of the national press in each country: 22% in Germany, 42% in France, 42% in Italy, 39% in Spain. As already mentioned, for each source, we can access articles belonging to one of the following 6 categories, according to the Dow Jones indexation: Commodity/Financial Market News (mcat), Corporate/Industrial News (ccat), Economic News (ecat), Politics/International Relations (gpir), Editorials (nedi), and Commentaries/Opinions (nedc). For each newspaper we get rid of duplicates before starting the analysis.

### Restricting news data to economic uncertainty news based on word embedding algorithms

The first step in creating our indexes is to gather all the news articles containing any form of the word “economy” and “uncertainty” (or corresponding translations of these two terms) based on the dictionary created by Azqueta-Gavaldón et al. (2023). These authors created associations of these terms for each country using word embedding algorithms which represent words in a high-dimensional vector space. In this vector space, each word is represented by a dense vector of real numbers, and semantically similar words are located

---

<sup>5</sup> In parenthesis we show the code of each category according to Dow Jones' indexation. Note, the EPU indexes constructed by Azqueta-Gavaldón et al. (2023) do not rely on this subset of press articles, but are based on the entire European press. Nevertheless, the resulting EPU indicators are similar, since the bulk of the articles related to economic uncertainty are contained in the aforementioned subset of press.

Table 1

**ECONOMY AND UNCERTAINTY-RELATED TERMS FOR EACH COUNTRY**

Country	Economy	Uncertainty
Germany	Wirtschaft, konjunktur, volkswirtschaft, okonomie	Unsicherheit, verunsicherung, ungewissheit
France	Économie, conjoncture	Incertitude, flou, inquiétude
Italy	Economia, congiunturali	Incertezza, instabilità, preoccupazione
Spain	Economía, económico	Incertidumbre, inquietud, desconfianza

**SOURCES:** Dow Jones and own calculations.

close together. This methodology creates a dictionary of terms that is specific to each country, and effectively takes into account the peculiarities of each particular language use. It also gives the user greater reassurance that the final results will be well-balanced across the geographies. The final terms used for each country are as follows (see Table 1).

The total number of news articles with related words to “economy” and “uncertainty” was 19,907 for Germany, 18,401 for France, 13,820 for Italy and 30,880 for Spain. Of these, 96% for Germany, 78% for France, 95% for Italy, and 96% for Spain, occurred after 2000.

### Topic modelling

Before feeding all the data (raw words per document) into the LDA algorithm to obtain unique topics, we need to pre-process them. *Stopwords*, punctuation, and numbers are removed.<sup>6</sup> All words are converted to lower case, and each word is converted to its root in a process known as “stemming”.<sup>7</sup>

Formally, within each article the probability of a word  $w_i$  is:

$$P(w_i) = \sum_{j=1}^k P(w_i | z_i = j) P(z_i = j) \quad (1)$$

where  $z_i$  is a latent variable indicating the topic from which the  $i$ th word was drawn and  $P(w_i | z_i = j)$  is the probability of word  $w_i$  being drawn from topic  $j$ .  $P(z_i = j)$  is the probability of drawing a word from topic  $j$  in the current article, which will vary across different articles. In other words,  $P(w | z)$  indicates which words are important to a topic, whereas  $P(z)$  is the frequency of those topics within an article. The algorithm maximises  $P(w_i | z_i = j)$  and  $P(z_i = j)$  from equation (1).

Direct maximisation turns out to be susceptible to problems of slow convergence or the algorithm getting stuck in local maxima (Griffiths and Steyvers, 2004). The two

<sup>6</sup> Stopwords are words that do not contain informative details about an article, e.g., “that” or “me”. Note that the list of stopwords is language-specific. We use the NLTK library, see [www.nltk.org/](http://www.nltk.org/).

<sup>7</sup> Stemming is language-specific and to carry it out, we use the SnowballStemmer: <https://www.nltk.org/modules/nltk/stem/snowball.html>.

most common methods used to approximate the posterior distribution given by these two probabilities are sampling methods (SM) and variational methods (VM). Although SM are asymptotically exact, they are very time consuming as they rely on techniques such as the Gibbs sampler. Alternatively, VMs approximate the posterior distribution of  $P(w_i | z_i = j)$  and  $P(w_i | z_i = j)$  using an alternative and simpler distribution:  $P(w | z)$ , and associated parameters. We use an advanced type of VM called *online variational Bayes* as proposed by Hoffman, Bach and Blei (2010) and the practical implementation of Rehurek and Sojka (2010). Finally, we unveil 30 topics for each country following Azqueta-Gavaldón et al. (2023).<sup>8</sup>

### Constructing the indicators

Following Azqueta-Gavaldón et al. (2023), we select those topics that best describe sources of policy uncertainty in the European context and that are also easy to identify across countries. These are the following 8 categories: fiscal; monetary; political; geopolitical; trade/manufacturing; European regulation; domestic regulation; and energy. As can be seen in Table 2, with the word stems that the LDA algorithm gives we can easily label each category/topic.<sup>9</sup> For example, the political topic is framed by words such as “ministry”, “president” or names of heads of state, while the monetary policy topic contains words such as “ECB”, “inflation” and “central bank” in their respective languages.

We construct the following indicators.

- **Topic-specific uncertainty components at the country level:** these are the topic proportions of the aforementioned 8 categories at the monthly level and for each country. These series are standardized to mean 100 and standard deviation of one in the reference period, which is defined as being the full time period available at the moment of writing the article (1997-2019 for DE, IT, ES, and 1996-2019 for FR). This makes them comparable across topics and countries.
- **EPU index at the country level:** this is computed as the sum by month of the topic proportions of the 8 topic-specific uncertainty components (before standardization). Then, the resulting sum is standardized to mean 100 and standard deviation of one in the reference period. This makes the EPU indexes comparable across countries.
- **Topic-specific uncertainty components for the euro area:** these indexes are computed as the weighted sum of the topic-specific uncertainty components for each country (standardized), using GDP in level as weight. Finally, the resulting indicator is standardized to mean 100 and standard deviation of one in the reference period.

<sup>8</sup> Using the log-likelihood score, Azqueta-Gavaldón et al. (2023) found 30 to 40 topics to be the optimal number of topics to unveil for each country. For additional information, see Azqueta-Gavaldón et al. (2023).

<sup>9</sup> Note that the words are “stemmed”, reduced to their word stem or base root, and hence, they may be difficult to read.

Table 2

**MOST RELEVANT WORD STEMS REPRESENTING GIVEN BY THE LDA FOR EACH COMPONENT**

	Germany Articles = 19,907	France Articles = 18,401	Italy Articles = 13,820	Spain Articles = 30,880
Monetary	ezb, notebank, geldpolit, fed, zentralbank, zins, inflation, zinerhoh, wahrung, mario, zinnsenk	taux, euro, banqu, monétaire, central, bce, économ, monnai, polit, européen, zon, fed	tass, bce, economy, banc, central, inflazion, monetar, eur, polit, drag, polit, merc	tip, bce, fed, inflacion, federal, monetari, reserv, central, interes, sub, trichet, banc, greensp
Fiscal	gesetz, unternehm, fall, hoh, heisst, regel, kunftig, deutsch, bundesregier, betroff, euro, moglich	gouvern, financ, milliard, fiscal, impôt, ministr, publicu, économ, plus, budget, dépens, mesur	fiscal, public, spes, tass, tagl, pag, cont, legg, redd, pension, impost, iva	deficit, impuest, fiscal, gast, pension, presupuest, reform, public, gobiern, presupuestari, recort, ingres
Political	spd, cdu, merkel, grun, csu, koalition, fdp, pertei, union, angela, scholz, seehof	président, ministr, premi, franc, gouvern, François, part, jean, polit, chef, plus, fait	president, ital, premier, palazz, trmont, parl, siciet, comon, democraz, elettoral, elezion, riform	pp, catalan, rajoy, ayuso, cataluny, zapater, sanchez, independent, eleccion, vot, candidat, electoral, inquierd
Geopolitical	rusland, russisch, putin, ukrain, sanktion, west, pol, ost, moskau, pen, osterreich, westlich	pay, part, militair, islam, immigr, musulman, russ, pay, président	russ, turc, putin, mosc, ucrain, sanzion, cines, gerr, usa, iran, global, american	rusi, iran, rus, guerra, israel, terror, islam, putin, putin, siria, militar, irak, arab, palestin
Trade /Manufacturing	europa, deutsch, deutschland, usa, polit, wirtschaft, land, prasident, unternehm, stark, wichtig, international	chin, américain, chinois, produit, prix, million, pétrol, prix, dollar	azi, merc, settor, export, industrial, fabbric, produzion, mad, marc	chin, eeuu, comercial, estadounidense, pekin, washington, ancel, mundial, indi, unid, comerci, guerr, omc, dolar
European Regulation	euro, land, griechenland, milliard, regier, iwf, staat, griechisch, geld, europa, zon, spani	européen, britann, unión, pay, ue, royaum, allemand, allemagn, merkel, berlin, angel, chanceli	europe, ital, eur, ue, grec, unión, econom, cris, german, merc, polit, franc,	europe, alemani, eu, itali, union, franci, brusel, país, comision, comunitari, britain, referendum, londr, aten, unidad
Domestic Regulation (a)	bank, deutsch, institute, kund, frankfurt, geschafft, kredit, chef, Commerzbank, geldhaus, geld, investmentbank	social, emploi, travail, banqu, financi, bancaire, sant, malad, médecin, det, milliard, plus	sindac, prod, confindustr, cgil, social, lavor, banc, cred, bancar, sistem, capitalm finanz	trabaj, empleo, laboral, contrat, sindicat, salari, par, ugt, desp, banc, deud, entidad, credit, bancari, million, financier
Energy	energi, erneuerbar, eon, energiew, strom, rwe, kraftwerk, teheran, eeg	nucléair, énerg, écolog edf, climat, électr gaz, central, énerget	ricerc, ambient, energ, digital, serviz, tav, ret, elettr, inform	energ, ciud, electr, gas, climat, energet, carbon, agu, metr, transport, emision

**SOURCES:** Dow Jones and own calculations.

**a** The topic Domestic regulation is composed of labour and domestic reforms as well as banking regulation.

- **EPU index for the euro area:** this is computed as a weighted sum of the country-specific EPU indexes (standardized), using GDP in level as weight. Finally, the resulting indicator is standardized to mean 100 and standard deviation of one in the reference period.

### 3 Economic policy uncertainty in the euro area

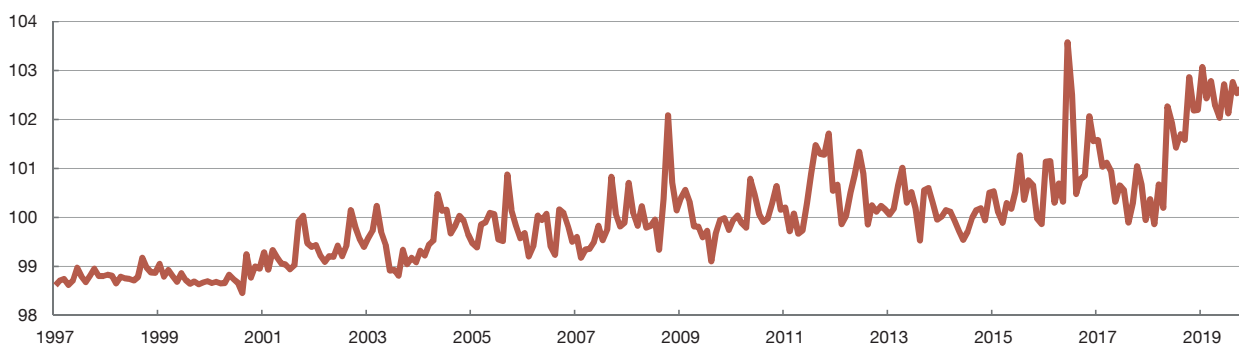
For each country, we compute the overall EPU index as the sum by month of the topic proportions of these 8 categories (topic-specific uncertainty components), which represent specific sources of economic policy uncertainty. Figures A1-A4 in the Appendix display the overall EPU index for Germany, France, Italy and Spain for the available time window (1997-2019 for DE, IT, ES, 1996-2019 for FR), along with the contribution to each topic-specific uncertainty components.<sup>10</sup>

Then, we construct the overall EPU index for the euro area as a weighted sum of the country-specific standardized EPU indicators, weighting by GDP. The latter is shown in Figure 1. According to this figure, the overall EPU indicator for the euro area experiences increases during times of economic distress. For example, it shows a significant increase during the Covid crisis, which had a profound impact on the European economy. Similarly, it increased significantly during the EU debt crisis, reflecting the high level of uncertainty surrounding economic policy decisions during that period.

One interesting observation from the figure is that the overall EPU for the euro area has remained higher after the great financial crisis, which occurred in 2008-2009. This is consistent with the findings of Baker et al. (2016), who showed that the level of economic policy uncertainty increased significantly after the crisis and has remained elevated since then. This suggests that the great financial crisis had a long-lasting impact on economic policy decisions and has contributed to a higher level of uncertainty in the years following the crisis.

Chart 1

#### EURO AREA EPU INDEX



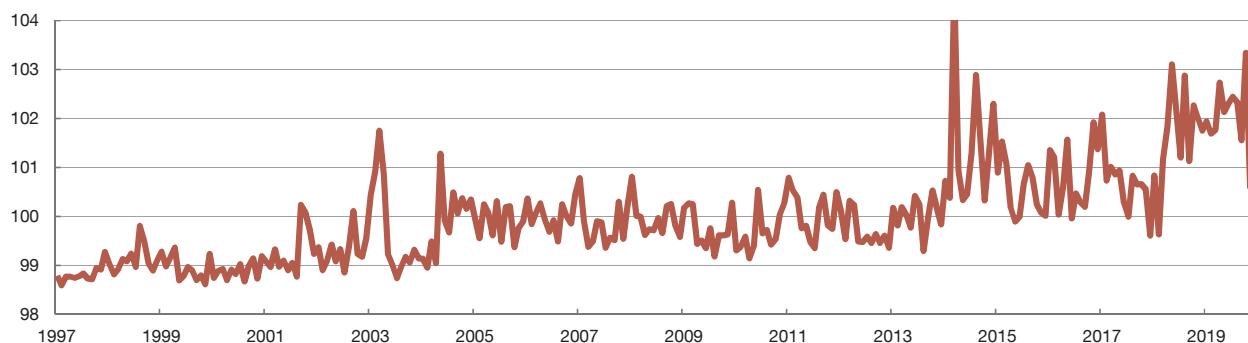
**SOURCES:** Dow Jones and own calculations.

**a** This index is standardized so that it has mean 100 and standard deviation of one.

<sup>10</sup> Note, for each country, the final overall EPU index is a standardized version of the aforementioned index displayed in Figures A1-A4 in the Appendix, in which we subtract the mean of the index and divide by its standard deviation, computed in the reference period, which is defined as being the full time period available at the moment of writing the article (1997-2019 for DE, IT, ES, and 1996-2019 for FR).

Chart 2

**EURO AREA GEOPOLITICAL UNCERTAINTY INDEX (GPU)**

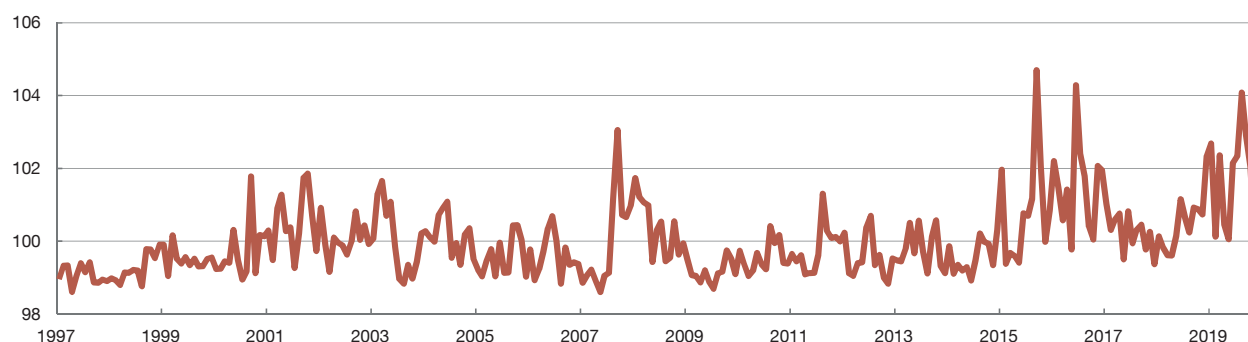


**SOURCES:** Dow Jones and own calculations.

**NOTE:** This index is standardized so that it has mean 100 and standard deviation of one.

Chart 3

**EURO AREA MONETARY POLICY UNCERTAINTY INDEX (MPU)**



**SOURCES:** Ghirelli et al. (2019), Baker et al. (2016), Azqueta-Gavaldón et al. (2023), Dow Jones and own calculations.

**NOTE:** This index is standardized so that it has mean 100 and standard deviation of one.

A value added of this methodology is that we can compute, for each country, indicators that are specific to the 8 sources of uncertainty that we have highlighted when constructing the index. As an example, the monetary policy component of euro area EPU (EPU-MPU) is based on the frequency of media references to monetary policy uncertainty in a given country, while the geopolitical component of euro area EPU (EPU-GPU) relies on the frequency of media references to geopolitical uncertainty. Then, we aggregate these topic-specific components across countries (weighting by GDP) to construct topic-specific EPU indicators for the Euro Area. These indicators reflect the evolution of specific components (sources) of economic policy uncertainty for the Euro Area. Fig. 2 and 3 display the geopolitical component of euro area EPU (EPU-GPU) and the monetary policy component of euro area EPU (EPU-MPU), respectively. As an example, the euro area EPU-MPU (Fig. 3) shows significant spikes during times of significant geopolitical unrest, such as the Iraq war in 2003,

the Arab Spring in 2011, the Russian invasion of Crimea in 2014, and the recent invasion of Ukraine (in the Spanish series that reaches this time period). These events generated significant uncertainty surrounding geopolitical developments and had a significant impact on global financial markets and the broader economy.

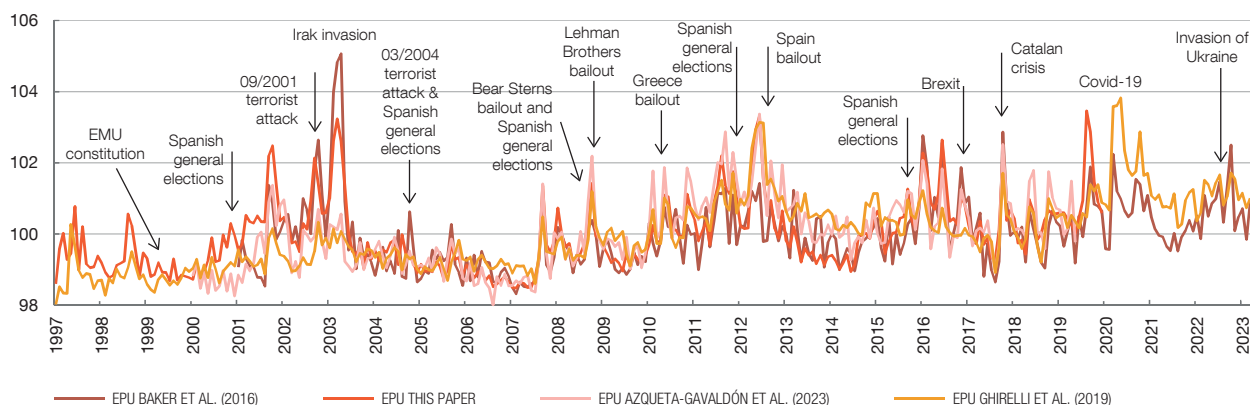
### The case of Spain: comparison across alternative EPU indexes

It is worth comparing our resulting EPU index for Spain with additional indices found in the literature. In this section we consider several additional EPU indicators available in the literature. The first one is the EPU indicator by Baker et al. (2016) which was first built based on a dictionary method and on the Spanish press. The second indicator that we compare ours with is Ghirelli et al. (2019) which is built by improving the list of keywords of the former to better capture Spanish policy uncertainty. Finally, we consider Azqueta-Gavaldón et al. (2023) as well, which is propose the topic modelling approach that we replicate in this paper.

As can be seen in Figure 4 all indicators behave very similarly. In all cases, the Spanish EPU index is shaped by some key events: i) the Iraq war at the beginning of 2003; ii) the sovereign debt crisis (December 2010) with its pinnacle during the period of banking recapitalization (June 2012); and iii) the Catalan independence referendum (October 2017); iv) the Covid 19 crisis (March 2020), and v) the Russian invasion of Ukraine (February 2022).

There are however, few differences among them worth discussing. The EPU by Baker et al. (2016) shows the highest peak around the geopolitical events of the war in Iraq in 2003. While this event was certainly a period with high EPU, it is hard to imagine that it was the period with the highest level among the whole time series. Besides this difference, we also find some differences in more recent period such as the Covid pandemic where the index of Ghirelli et al. (2019) displays a higher peak than the rest of indicators (excluding

Chart 4  
EPU NARRATIVE FOR SPAIN



SOURCE: XXX.

NOTE: For comparability, each EPU index is standardized so to have mean 100 and standard deviation of one in the considered period. Time span considered: EPU Ghirelli et al. (2019): Jan1997-Mar2023; EPU Baker et al. (2016): Jan2001-Mar2023; EPU Azqueta-Gavaldón et al. (2023): Jan2000-Dic2019; EPU this paper: Jan1997-Dic2019.



Table 3

**CORRELATION MATRIX BETWEEN ALTERNATIVE EPU INDICATORS FOR SPAIN**

	EPU: this paper	EPU Azqueta-Gavaldón et al. (2023) EPU	EPU Baker et al. (2016)	EPU Ghirelli et al. (2019)
EPU: this paper	1			
EPU Azqueta-Gavaldón et al. (2023)	0.74	1		
EPU Baker et al. (2016)	0.78	0.60	1	
EPU Ghirelli et al. (2019)	0.60	0.83	0.53	1

**SOURCES:** Ghirelli et al. (2019), Baker et al. (2016), Azqueta-Gavaldón et al. (2023), Dow Jones and own calculations.

**NOTE:** Pair-wise correlations are computed considering all sample available, that is: for EPU Ghirelli et al. (2019): Jan1997-Mar2023; for EPU Baker et al. (2016): Jan2001-Mar2023; for EPU Azqueta-Gavaldón et al. (2023): Jan2000-Dic2019; for EPU this paper: Jan1997-Dic2019.

Azqueta-Gavaldón et al. (2023) which terminates in 2019). The high level of EPU around this time is coherent with the fact that there was an extreme uncertainty around lockdown, subsidies, and travelling policies. All in all, these indices display really similar dynamics and high correlations (see Table 3).

## 4 Conclusions

In this paper, we have presented an updated and more nuanced approach to measuring Economic Policy Uncertainty (EPU) using Latent Dirichlet Allocation (LDA). Specifically, we adopted the method described in Azqueta-Gavaldón et al. (2023), which identified eight distinct components of EPU tailored to the European case. Despite the differences in data and time span considered, our results are strongly in line with the original paper, which suggests that the method is highly versatile and can be applied in various contexts. By providing a more detailed and granular understanding of the sources of EPU, we believe that our approach can offer valuable insights to policymakers and researchers interested in understanding the impact of policy uncertainty on economic outcomes. Our findings highlight the importance of considering the underlying topics and themes that contribute to economic policy uncertainty, and we hope that our work will encourage further research in this area.

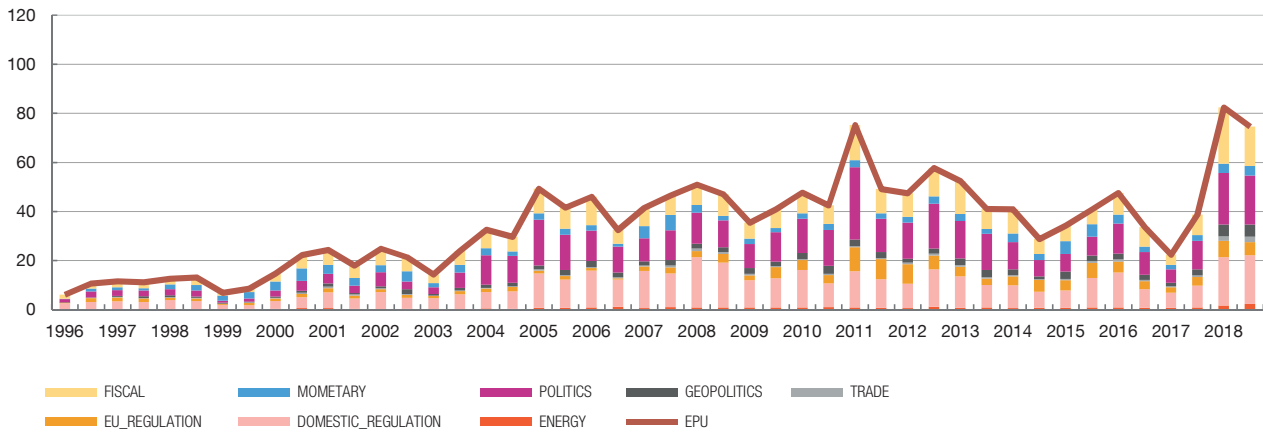
## References

- Andres-Escayola, Erik, Corinna Ghirelli, Luis Molina, Javier J. Pérez and Elena Vidal. (2022). "Using newspapers for textual indicators: which and how many?". Documentos de Trabajo, 2235, Banco de España. <https://www.bde.es/ff/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadadas/DocumentosTrabajo/22/Files/dt2235e.pdf>
- Azqueta-Gavaldón, Andrés. (2017). "Developing news-based economic policy uncertainty index with unsupervised machine learning". *Economics Letters*, 158, pp. 47-50. <https://doi.org/10.1016/j.econlet.2017.06.032>
- Azqueta-Gavaldón, Andrés, Dominik Hirschbühl, Luca Onorante and Lorena Saiz. (2023). "Sources of Economic Policy Uncertainty in the euro area". *European Economic Review*, 152(104373). <https://doi.org/10.1016/j.euroecorev.2023.104373>
- Baker, Scott R., Nicholas Bloom and Steven J. Davis. (2016). "Measuring economic policy uncertainty." *The Quarterly Journal of Economics*, 131(4), pp. 1593-1636. <https://doi.org/10.1093/qje/qjw024>
- Ghirelli, Corinna, Javier J. Pérez and Alberto Urtasun. (2019). "A new economic policy uncertainty index for Spain". *Economics Letters*, 182, pp. 64-67. <https://doi.org/10.1016/j.econlet.2019.05.021>
- Griffiths, Thomas L., and Mark Steyvers. (2004). "Finding scientific topics". *Proceedings of the National academy of Sciences*, 101(suppl. 1), pp. 5228-5235. <https://doi.org/10.1073/pnas.0307752101>
- Hoffman, Matthew, Francis Bach and David Blei. (2010). "Online learning for Latent Dirichlet Allocation." In Lafferty, John D., Christopher K. I. Williams, John Shawe-Taylor, Richard S. Zemel and Aron Culotta (eds.), *Advances in Neural Information Processing Systems 23: 24th Annual Conference on Neural Information Processing Systems 2010. Proceedings of a meeting held 6-9 December 2010, Vancouver, British Columbia, Canada*. Curran Associates, Inc. [https://papers.nips.cc/paper\\_files/paper/2010/file/71f6278d140af599e06ad9bf1ba03cb0-Paper.pdf](https://papers.nips.cc/paper_files/paper/2010/file/71f6278d140af599e06ad9bf1ba03cb0-Paper.pdf)
- Kaveh-Yazdy, Fatemeh, and Sajjad Zarifzadeh. (2021). "Measuring economic policy uncertainty using an unsupervised word embedding-based method". ArXiv preprint arXiv:2105.04631. <https://doi.org/10.48550/arXiv.2105.04631>
- Naboka-Krell, Viktoriia. (2023). "Construction and Analysis of Uncertainty Indices based on Multilingual Text Representations". MAGKS Papers on Economics, 202310, Philipps-Universität Marburg, Faculty of Business Administration and Economics, Department of Economics. <https://EconPapers.repec.org/RePEc:mar:magkse:202310>
- Řehůřek, Radim, and Petr Sojka. (2010). "Software framework for topic modelling with large corpora". *Conference: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. <https://doi.org/10.13140/2.1.2393.1847>

## Appendix

Chart A.1

### EPU INDEX BY TOPIC-RELATED COMPONENTS: ITALY

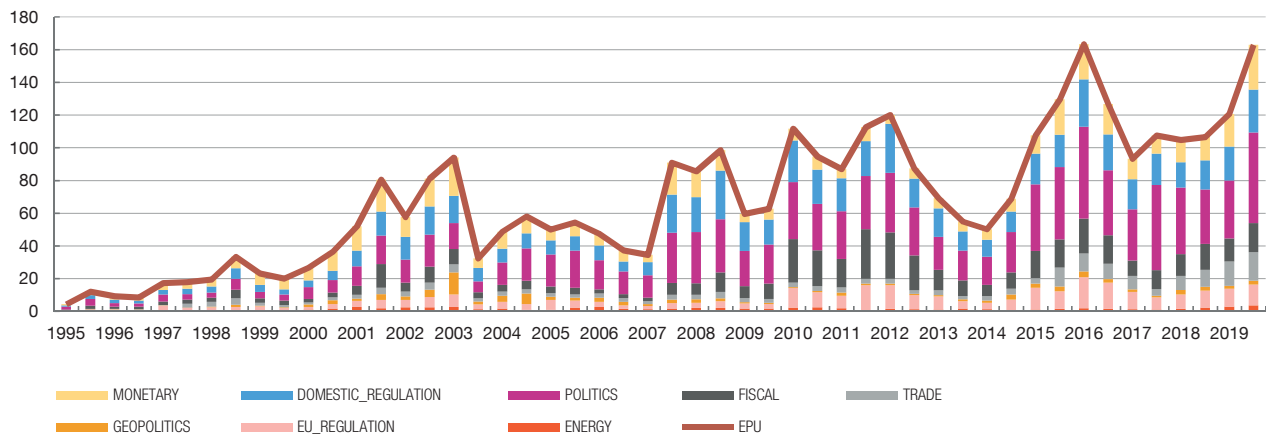


**SOURCES:** Dow Jones and own calculations.

**NOTES:** The EPU index is not standardized: each monthly value is the sum of the topic-specific uncertainty components.

Chart A.2

### EPU INDEX BY TOPIC-RELATED COMPONENTS: SPAIN

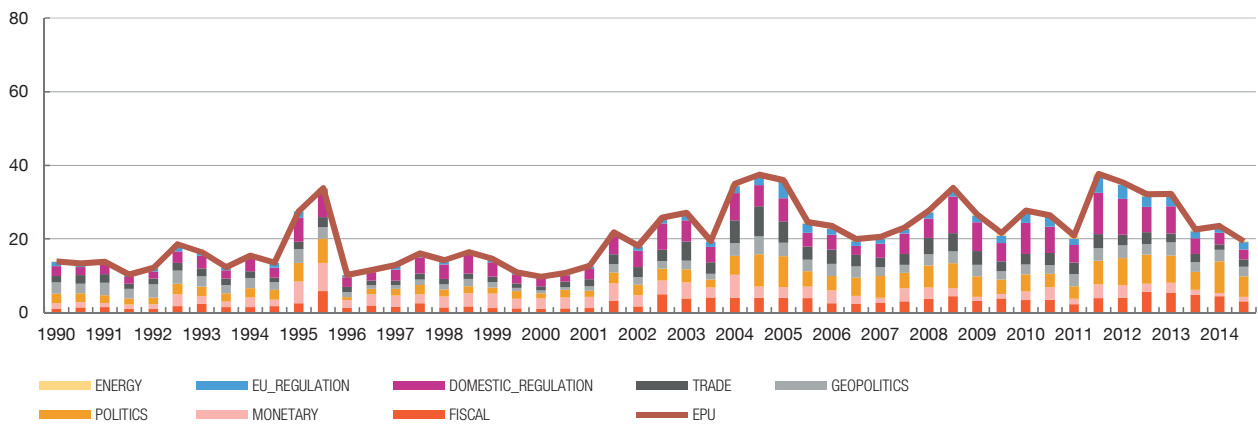


**SOURCES:** Dow Jones and own calculations.

**NOTES:** The EPU index is not standardized: each monthly value is the sum of the topic-specific uncertainty components.

Chart A.3

**EPU INDEX BY TOPIC-RELATED COMPONENTS: FRANCE**

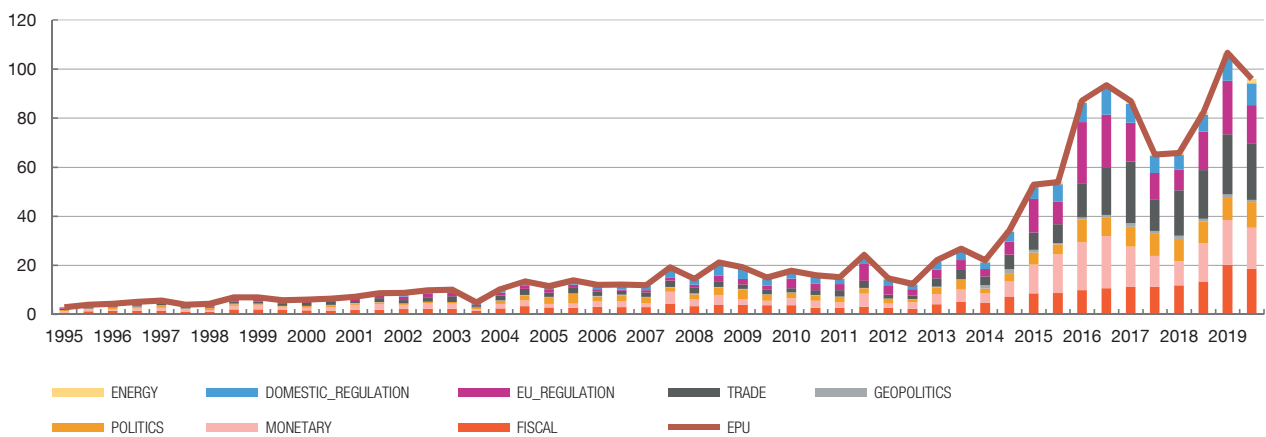


**SOURCES:** Dow Jones and own calculations.

**NOTES:** The EPU index is not standardized: each monthly value is the sum of the topic-specific uncertainty components.

Chart A.4

**EPU INDEX BY TOPIC-RELATED COMPONENTS: GERMANY**



**SOURCES:** Dow Jones and own calculations.

**NOTE:** The EPU index is not standardized: each monthly value is the sum of the topic-specific uncertainty components.

## BANCO DE ESPAÑA PUBLICATIONS

### OCCASIONAL PAPERS

- 2130 PABLO HERNÁNDEZ DE COS: Testimony before the Congress of Deputies Budget Committee on 25 October 2021 and before the Senate Budget Committee on 30 November 2021 in relation to the Draft State Budget for 2022. (There is a Spanish version of this edition with the same number).
- 2131 LAURA AURIA, MARKUS BINGMER, CARLOS MATEO CAICEDO GRACIANO, CLÉMENCE CHARAVEL, SERGIO GAVILÁ, ALESSANDRA IANNAMORELLI, AVIRAM LEVY, ALFREDO MALDONADO, FLORIAN RESCH, ANNA MARIA ROSSI and STEPHAN SAUER: Overview of central banks' in-house credit assessment systems in the euro area.
- 2132 JORGE E. GALÁN: CREWS: a CAMELS-based early warning system of systemic risk in the banking sector.
- 2133 ALEJANDRO FERNÁNDEZ CERESO and JOSÉ MANUEL MONTERO: A sectoral analysis of the future challenges facing the Spanish economy. (There is a Spanish version of this edition with the same number).
- 2201 MANUEL A. PÉREZ ÁLVAREZ: New allocation of special drawing rights. (There is a Spanish version of this edition with the same number).
- 2202 PILUCA ALVARGONZÁLEZ, MARINA GÓMEZ, CARMEN MARTÍNEZ-CARRASCAL, MYROSLAV PIDKUYKO and ERNESTO VILLANUEVA: Analysis of labor flows and consumption in Spain during COVID-19.
- 2203 MATÍAS LAMAS and SARA ROMANIEGA: Designing a price index for the Spanish commercial real estate market. (There is a Spanish version of this edition with the same number).
- 2204 ÁNGEL IVÁN MORENO BERNAL and TERESA CAMINERO GARCÍA: Analysis of ESG disclosures in Pillar 3 reports. A text mining approach.
- 2205 OLYMPIA BOVER, LAURA CRESPO and SANDRA GARCÍA-URIBE: Household indebtedness according to the Spanish Survey of Household Finances and the Central Credit Register: a comparative analysis. (There is a Spanish version of this edition with the same number).
- 2206 EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO and ROBERTO RAMOS: Population dynamics during the COVID-19 pandemic. (There is a Spanish version of this edition with the same number).
- 2207 JULIO GÁLVEZ: Measuring the equity risk premium with dividend discount models.
- 2208 PILAR CUADRADO, MARIO IZQUIERDO, JOSÉ MANUEL MONTERO, ENRIQUE MORAL-BENITO and JAVIER QUINTANA: The potential growth of the Spanish economy after the pandemic. (There is a Spanish version of this edition with the same number).
- 2209 PANA ALVES, SERGIO MAYORDOMO and MANUEL RUIZ-GARCÍA: Corporate financing in fixed-income markets: the contribution of monetary policy to lowering the size barrier. (There is a Spanish version of this edition with the same number).
- 2210 PABLO BURRIEL, IVÁN KATARYNIUK and JAVIER J. PÉREZ: Computing the EU's SURE interest savings using an extended debt sustainability assessment tool.
- 2211 LAURA ÁLVAREZ, ALBERTO FUERTES, LUIS MOLINA and EMILIO MUÑOZ DE LA PEÑA: Fund raising in the international capital markets in 2021. (There is a Spanish version of this edition with the same number).
- 2212 CARLOS SANZ: El peso del sector público en la economía: resumen de la literatura y aplicación al caso español.
- 2213 LEONOR DORMIDO, ISABEL GARRIDO, PILAR L'HOTELLERIE-FALLOIS and JAVIER SANTILLÁN: Climate change and sustainable growth: international initiatives and European policies. (There is a Spanish version of this edition with the same number).
- 2214 CARMEN SÁNCHEZ and JARA QUINTANERO: Las empresas *finotech*: panorama, retos e iniciativas.
- 2215 MARÍA ALONSO, EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO, DIANA POSADA, PATROCINIO TELLO-CASAS and CARLOS TRUCHARTE: In-person access to banking services in Spain: a comparison with other countries and other types of services. (There is a Spanish version of this edition with the same number).
- 2216 BEATRIZ GONZÁLEZ, ENRIQUE MORAL-BENITO and ISABEL SOLER: Schumpeter Meets Goldilocks: the Scarring Effects of Firm Destruction.
- 2217 MARIO ALLOZA, JÚLIA BRUNET, VICTOR FORTE-CAMPOS, ENRIQUE MORAL-BENITO and JAVIER J. PÉREZ: Government spending in Spain from a European perspective. (There is a Spanish version of this edition with the same number).
- 2218 PABLO AGUILAR, BEATRIZ GONZÁLEZ and SAMUEL HURTADO: Carbon tax sectoral (CATS) model: a sectoral model for energy transition stress test scenarios.
- 2219 ALEJANDRO MUÑOZ-JULVE and ROBERTO RAMOS: Estimation of the impact of changes in the period used to calculate the regulatory base on new retirement pension amounts. (There is a Spanish version of this edition with the same number).

- 2220 LUIS ÁNGEL MAZA: An estimation of the carbon footprint in Spanish credit institutions' business lending portfolio. (There is a Spanish version of this edition with the same number).
- 2221 SUSANA MORENO SÁNCHEZ: The EU-UK relationship: regulatory divergence and the level playing field.
- 2222 ANDRÉS ALONSO-ROBISCO and JOSÉ MANUEL CARBÓ: Inteligencia artificial y finanzas: una alianza estratégica.
- 2223 LUIS FERNÁNDEZ LAFUERZA, MATÍAS LAMAS, JAVIER MENCÍA, IRENE PABLOS and RAQUEL VEGAS: Analysis of the usability of capital buffers during the crisis precipitated by COVID-19. (There is a Spanish version of this edition with the same number).
- 2224 SONSOLES GALLEGO, ISABEL GARRIDO and IGNACIO HERNANDO: IMF precautionary facilities and their use in Latin America. (There is a Spanish version of this edition with the same number).
- 2301 LAURA HOSPIDO, CARLOS SANZ and ERNESTO VILLANUEVA: Air pollution: a review of its economic effects and policies to mitigate them.
- 2302 IRENE MONASTEROLO, MARÍA J. NIETO and EDO SCHETS: The good, the bad and the hot house world: conceptual underpinnings of the NGFS scenarios and suggestions for improvement.
- 2303 ADRIÁN LÓPEZ GONZÁLEZ: Inteligencia artificial aplicada al control de calidad en la producción de billetes.
- 2304 BELÉN AROCA MOYA: Conceptos, fundamentos y herramientas de neurociencia, y su aplicación al billete.
- 2305 MARÍA ALONSO, EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO, DIANA POSADA and PATROCINIO TELLO-CASAS: Un repaso de las diversas iniciativas desplegadas a nivel nacional e internacional para hacer frente a los riesgos de exclusión financiera.
- 2306 JOSÉ LUIS ROMERO UGARTE, ABEL SÁNCHEZ MARTÍN and CARLOS MARTÍN RODRÍGUEZ: Alternativas a la evolución de la operativa bancaria mayorista en el Eurosistema.
- 2307 HENRIQUE S. BASSO, OURANIA DIMAKOU and MYROSLAV PIDKUYKO: How inflation varies across Spanish households.
- 2308 LAURA CRESPO, NAJIBA EL AMRANI, CARLOS GENTO and ERNESTO VILLANUEVA: Heterogeneidad en el uso de los medios de pago y la banca *online*: un análisis a partir de la Encuesta Financiera de las Familias (2002-2020).
- 2309 HENRIQUE S. BASSO, OURANIA DIMAKOU and MYROSLAV PIDKUYKO: How consumption carbon emission intensity varies across Spanish households.
- 2310 IVÁN AUCIELLO-ESTÉVEZ, JOSEP PIJOAN-MAS, PAU ROLDAN-BLANCO and FEDERICO TAGLIATI: Dual labor markets in Spain: a firm-side perspective.
- 2311 CARLOS PÉREZ MONTES, JORGE E. GALÁN, MARÍA BRU, JULIO GÁLVEZ, ALBERTO GARCÍA, CARLOS GONZÁLEZ, SAMUEL HURTADO, NADIA LAVÍN, EDUARDO PÉREZ ASENJO and IRENE ROIBÁS: Marco de análisis sistémico del impacto de los riesgos económicos y financieros.
- 2312 SERGIO MAYORDOMO and IRENE ROIBÁS: La traslación de los tipos de interés de mercado a los tipos de interés bancarios.
- 2313 CARLOS PÉREZ MONTES, ALEJANDRO FERRER, LAURA ÁLVAREZ ROMÁN, HENRIQUE BASSO, BEATRIZ GONZÁLEZ LÓPEZ, GABRIEL JIMÉNEZ, PEDRO JAVIER MARTÍNEZ-VALERO, SERGIO MAYORDOMO, ÁLVARO MENÉNDEZ PUJADAS, LOLA MORALES, MYROSLAV PIDKUYKO and ÁNGEL VALENTÍN: Marco de análisis individual y sectorial del impacto de los riesgos económicos y financieros.
- 2314 PANA ALVES, CARMEN BROTO, MARÍA GIL and MATÍAS LAMAS: Indicadores de riesgos y vulnerabilidades en el mercado de la vivienda en España.
- 2215 ANDRÉS AZQUETA-GAVALDÓN, MARINA DIAKONOVA, CORINNA GHIRELLI and JAVIER J. PÉREZ: Sources of economic policy uncertainty in the euro area: a ready-to-use database.