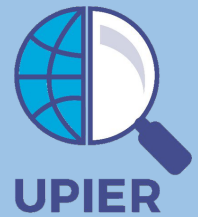


---

# USES OF THE PAST IN INTERNATIONAL ECONOMIC RELATIONS

---



## MINING DIE ZEIT: A STRUCTURAL TOPIC MODEL ANALYSIS OF SPAIN'S IMAGE IN THE GERMAN MEDIA 1946-2009

Anselm Küsters and Elisa Garrido Moreno

**UPIER WORKING  
PAPER Vol. 16**

**SEPTEMBER 2019**

## Foreword

UPIER Working Papers series reflects the work in progress of the researchers associated with the HERA funded project Uses of the Past in International Economic Relations and of others whose papers directly address UPIER research themes. The papers are peer reviewed by UPIER and associated researchers and seek to advance our understanding of how the past has been constructed and used in international economic relations over the past 200 years.

The views expressed in this working paper, and all errors and omissions, should be regarded as those solely of the authors and are not necessarily the views of the affiliated institutions.

For more information on UPIER visit: <https://upier.web.ox.ac.uk/>

## Authors

Anselm Küsters is a PhD. candidate at the Max Planck Institute for European Legal History, Frankfurt am Main. He holds a MPhil in Economic and Social History from Oxford University. His doctoral research focuses on the impact of German ordoliberalism on EU competition policy (<https://www.rg.mpg.de/kuesters/en>). Contact email address: [kuesters@rg.mpg.de](mailto:kuesters@rg.mpg.de)

Elisa Garrido Moreno is a Postdoctoral Fellow at University Carlos III Madrid. Since September 2019, she will be a Visiting Research Fellow at the Ruskin School of Art, Oxford. She holds a PhD in Art Theory and History. Her research interests include also cultural history, visual culture and digital humanities (<http://uc3m.academia.edu/ElisaGarrido>). Contact email address: [elisa.garrido@uc3m.es](mailto:elisa.garrido@uc3m.es)

**ABSTRACT**

The onset of the Eurozone crisis popularised the PIGS acronym; it was during this period that it came to convey the allegedly 'lax' Southern European attitudes towards inflation and state deficit, and contrasted Portugal, Italy, Greece, and Spain unfavourably with their Northern European neighbours. In order to investigate the historic formation of this country group heuristic we apply a Structural Topic model (STM) to all 2,443 articles published in the German newspaper *Die Zeit* that mention Spain between 1946 and 2009. How the topics covered in these articles evolved over time shows that a close association between Spain with other Southern European countries was an enduring characteristic of the German media discourse over a period of more than sixty years. Moreover, STM allows us to distinguish and characterise the cultural, political and an economic dimensions of the way PIGS countries were portrayed in the German media. We find that the initial PIGS heuristic, shaped by socio-cultural attitudes that mainly reflected the experiences of German tourists, was later supplemented with economic ascriptions, which in turn provided the basis for the moralising media discourse accompanying the Eurozone crisis.

## INTRODUCTION

Since the onset of the so-called Eurozone crisis in 2008, member states have faced several problems that challenge the viability of the monetary union, most importantly undercapitalized banks, rising bond yields (a marked problem in Southern countries), and slowing growth rates.<sup>1</sup> Throughout this period, academics and commentators alike have aimed to explain these interlocking crises as well as their divergent perceptions in different European countries. It has been argued that we are witnessing the ongoing expression of cultural differences that originated in the complex late sixteenth-century, with its conflict between the Protestant North and the Catholic South, and the development of multiple socio-cultural factors that tended to pull the two types of societies apart.<sup>2</sup> Such historical divergences in religious and cultural values have been identified repeatedly as the source of the dilemma between European nations, and their differing responses to the Eurozone crisis, which has shaped the past decade of European politics.<sup>3</sup> Others have argued that this crisis originated in the different economic

---

<sup>1</sup> Shambaugh, J., 'The Euro's Three Crises', *Brookings Papers on Economic Activity*, 43/1 (2012), pp. 157-231.

<sup>2</sup> Hien, J., 'The Religious Foundations of the European Crisis', *JCMS: Journal of Common Market Studies*, 57/2 (2019), pp. 185-204; Akaliyski, P., 'United in diversity? The convergence of cultural values among EU member states and candidates', *European Journal of Political Research*, 58/2 (2018), pp. 388-411. See also the foundational study by Elliott, J., *Europe Divided* (London, 2000), pp. 1559-1598.

<sup>3</sup> It has to be noted that authors like Dainotto have challenged those who think that Europe is only the result of the imposition of Northern thinking by reviewing those historical theories about South-North differences from Montesquieu to the present. Dainotto, R., *Europe. In Theory* (Durham, 2007).

philosophies of Germany and France, and that the North South divide is therefore an extrapolation of the 'Rhine-divide' in intellectual history.<sup>4</sup>

However, there were other moments of encounter when Northern and the Southern Europe became intermingled, and this points to a more complex story about the way the two sides have perceived and influenced each other over time. We need to remember that different Northern European countries called for workers from the Mediterranean Basin to meet acute labour shortages, and that between 1946 and 2009 around five million people migrated from the Mediterranean area to Northern Europe<sup>5</sup> Attitudes would shift again once migration renewed, following the economic crisis of 2007-2008, when rising unemployment rates, especially among young people,<sup>6</sup> and subsequent austerity measures provoked thousands of Portuguese, Italian, Greek and Spanish to leave the regions that had been hit hardest by the crisis.<sup>7</sup>

This paper focuses on a media phenomenon that occurred when financial crisis prompted the historical socio-cultural differences between North and South to be re-expressed in economic terms through the increased popularity of the acronym PIGS. PIGS came to refer to EU members states that struggled to refinance their government debt during the crisis

(that is, Portugal, Italy, Greece, and Spain), perpetuating the stereotype of two Europes. Seeing Southern European countries as a cluster, or grouping, would become increasingly problematic once 'PIGS' began to reference their peculiar economic histories with regards to inflation and state deficit in discussions on European Monetary Union.<sup>8</sup> The subsequent Eurozone crisis influenced the ways the press spoke about PIGS members: it popularised images and ideas that in turn influenced political choices.<sup>9</sup> The socio-

---

<sup>4</sup> Brunnermeier, M.K. et al., *The Euro and the Battle of Ideas* (Princeton, 2016).

<sup>5</sup> Zimmerman, K.F., 'European Migration: Push and Pull', *International Regional Science Review*, 19/1-2 (1996), pp. 95-128.

<sup>6</sup> See Table 3 in: Zamora-Kapoor, A., and Coller, X., 'The Effects of the Crisis: Why Southern Europe?', *American Behavioral Scientist*, 58/12 (2014), pp. 1511-1516.

<sup>7</sup> Lafleur, J., and Stanek, M. (eds.), *South-North Migration of EU Citizens in Times of Crisis* (Dordrecht, 2017); Glorius, B. and Domínguez-Mujica, J., *European Mobility in Times of Crisis: The New Context of European South-North Migration* (Bielefeld, 2017).

<sup>8</sup> Gros, D., 'How Fit are the Candidates for EMU?', *The World Economy*, 23/10 (2000), pp. 1367-1377; Mundell, R., 'Currency Areas, Common Currencies and EMU', *The American Economic Review*, 87/2 (1997), pp. 214-216.

<sup>9</sup> Blyth, M., 'Structures Do Not Come with an Instruction Sheet: Interests, Ideas, and Progress in Political Science', *Perspectives on Politics*, 1/4 (2003), pp. 695-706.

economic reasons for grouping these countries together in debates on ‘the Southern problem’ appear to have shaped market participants’ perceptions exacerbating the economic differences within Europe.<sup>10</sup>

In this paper we aim to further this debate by looking at the narrative construction of financial crisis through the use of the acronym PIGS. To do so, we analyse all 2,443 articles published between 1946 and 2009 in the German newspaper *Die Zeit* that mention Spain as a keyword. Since previous research suggests that most articles published under the topic ‘Spain’ in *Die Zeit* also referred to Portugal, Italy and Greece,<sup>11</sup> the emergence of a country group heuristic can be analysed through correlations between Spain and the PIGS countries’ co-occurring names.

These heuristics can be empirically observed and analysed through a text mining approach, which uses traditional primary sources – such as newspaper articles, social media, or survey data – as inputs for a variety of recently developed statistical approaches that offer novel and quantitative insights into the underlying semantic data. At the forefront of this development are so-called topic models: algorithms that allow for a machine-assisted reading of large collections of documents (known as ‘corpora’) in order to automatically infer content from them.<sup>12</sup> Topic modelling has been successfully employed in recent historical enquiries related to newspapers:<sup>13</sup> it has been used to identify themes in a colonial US newspaper, the *Pennsylvania Gazette*, published between 1728 and 1800;<sup>14</sup> to measure the influence of varying terror alert levels under the Bush administration on the public discourse;<sup>15</sup> and to trace the

---

<sup>10</sup> The PIGS heuristic had a clear negative effect on the PIGS countries’ market treatment: Brazys, S. and Hardiman, N., ‘The ‘PIIGS’ acronym had a clear negative impact on the response of financial markets to the “PIIGS countries” during the crisis, *LSE Blog Post* (12.12.2014), <https://blogs.lse.ac.uk/euoppblog/2014/12/12/the-piigs-acronym-had-a-clear-negative-impact-on-the-market-treatment-of-the-piigs-countries-during-the-crisis/> (accessed: 29.08.2019).

<sup>11</sup> See Table 1 in: Garrido, E., ‘Exploring North-South Identities Using NLP: The Image of Spain in the German Weekly *Die Zeit*’, *Bulletin for Spanish and Portuguese Historical Studies*, 43/1 (2018), pp. 53-69.

<sup>12</sup> This methodology was initiated by the development of probabilistic topic models, such as Latent Dirichlet Allocation: Blei, D.M., Ng, A., Jordan, M., ‘Latent Dirichlet Allocation’, *Journal of Machine Learning Research*, 3 (2003), pp. 993-1022. These models were subsequently extended, see e.g. the Correlated Topic model: Blei, D.M. and Lafferty, J.D., ‘A correlated topic model of Science’, *The Annals of Applied Statistics*, 1/1 (2007), pp. 17-35.

<sup>13</sup> See also the general literature overview of Topic modeling applications in: Wehrheim, L., ‘Economic history goes digital: topic modeling the *Journal of Economic History*’, *Cliometrica* 13 (2019), pp. 83-125.

<sup>14</sup> Newman D.J. and Block, S., ‘Probabilistic topic decomposition of an eighteen-century American newspaper’, *J Am Soc Inform Sci Technol*, 57/6 (2006), pp. 753-767.

<sup>15</sup> Bonilla, T. and Grimmer, J., ‘Elevated threat levels and decreased expectations: how democracy handles terrorist threats’, *Poetics*, 41/6 (2013), pp. 650-669.

media coverage of public support of the arts in the US.<sup>16</sup> Topic modelling has also proved to be a particularly useful tool for historical research when dealing with thousands of newspaper articles over a long period of time, such as the recent analysis of the coverage of nuclear technology in the *New York Times* between 1945 and 2013.<sup>17</sup>

In this paper, we apply a specific type of topic model, a so-called structural topic model (STM), to our corpus of *Zeit* articles.<sup>18</sup> STM has been recently used to model the framing of international newspapers,<sup>19</sup> as well as other forms of modern media communication such as Twitter feeds,<sup>20</sup> but to the best of our knowledge, this paper is the first one to apply STM to articles published in *Die Zeit*. In contrast to standard topic modelling exercises, STM permits users to incorporate specific metadata (e.g. author name, or year of publication) into the model. This allows us to define topics and estimate their relationship to the metadata and to conduct hypothesis-testing about these relationships. Here, we aim to empirically check the hypothesis, outlined in detail in UPIER Working Paper Vol 8,<sup>21</sup> that the strong association of Spain with other Southern European countries has been an ongoing characteristic of German discourse in the media for more than sixty years and that this association has cultural, political and economic dimensions. All three discourses (cultural, political, economic) regarding PIGS countries can be captured empirically by the estimated STM presented in this paper. We find that the initial PIGS heuristic, shaped by socio-cultural attitudes that mainly reflected the experiences of German tourists, was later supplemented with economic ascriptions. In turn, this provided the basis for the moralising media discourse during the Eurozone crisis. In this approach, we hope to provide a bridge between qualitative research goals and statistical techniques that might provide inspiration for other social scientists as well.

---

<sup>16</sup> DiMaggio, P., Nag, M. and Blei, D., 'Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. Government arts funding', *Poetics*, 41/6 (2013), pp. 570-606.

<sup>17</sup> Jacobi, C., van Atteveldt, W. and Welbers, K., 'Quantitative analysis of large amounts of journalistic texts using topic modeling', *Digit J*, 4/1 (2015), pp. 89-106.

<sup>18</sup> This is done via the R language *stm* package.

<sup>19</sup> Roberts, M.E., Stewart, B.M. and Airoldi, E., 'A model of text for experimentation in the social sciences', *Journal of the American Statistical Association*, 111/515 (2016), pp. 988-1003.

<sup>20</sup> Lucas, C., Nielsen, R., Roberts, M., Stewart, B., Storer, A., Tingley, D., 'Computer Assisted Text Analysis for Comparative Politics', *Political Analysis*, 23/2 (2015), pp. 254-277.

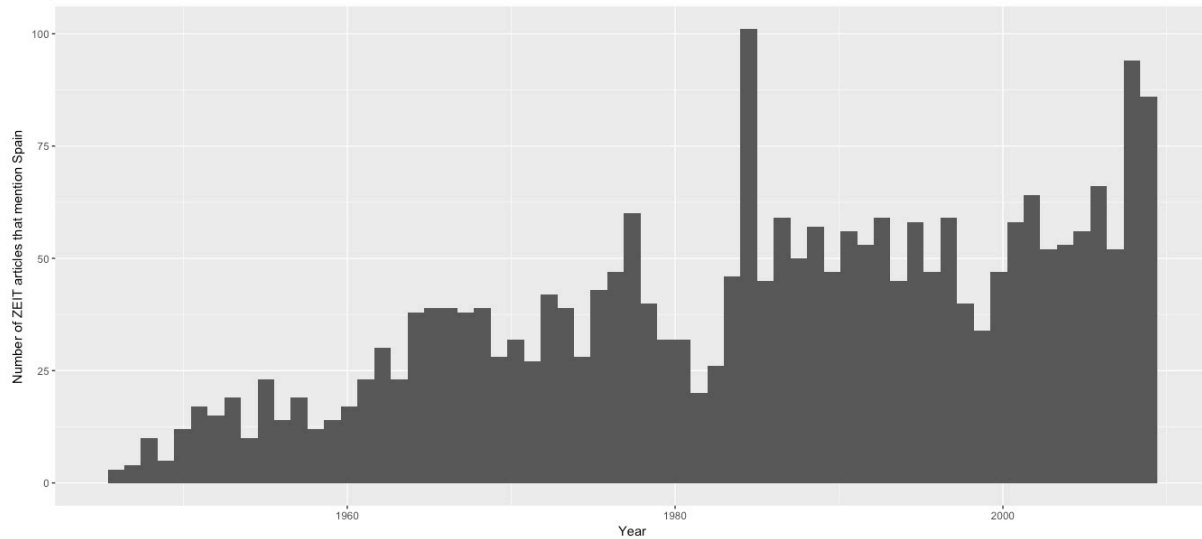
<sup>21</sup> Garrido, E., 'Exploring North-South Identities Using NLP: The Image of Spain in the German Weekly *Die Zeit*', *Bulletin for Spanish and Portuguese Historical Studies*, 43/1 (2018), pp. 53-69.

The outline of this paper is as follows. In section two we introduce the new dataset on Spain's image in the German press, including descriptive statistics and word correlations that allow us to characterise the underlying discourses related to PIGS countries. In section three we provide an introduction to STM and details of its specific implementation followed in this paper. Sections four and five cover our main results related to STM, accompanied by several visualisations that allow for an understanding of the temporal dimension of the cultural, political and economic discourses regarding PIGS countries as expressed in *Zeit* articles. We draw conclusions in section six.

### **THE ZEIT DATASET: CONSTRUCTION AND DESCRIPTIVE STATISTICS**

To conduct text mining analysis, including STM quantitative analysis, we use the whole population of *Zeit* newspaper articles about Spain that were written between 1946 and 2009 (2,443 articles in total), thus covering the whole process of post-war reconstruction, the Cold War, European integration, and the first years of the Great Recession. The articles were gathered via the open source software *DiaCollo*, which allows one to search through the *Deutsches Textarchiv*, a large collection of German texts published in different fields between 1650 and 1900, including the entire collection of articles published in *Die Zeit* (1946-2009). The distribution of individual articles within this corpus is plotted over time in the figure below. As can be seen, there has been a steady increase in *Zeit* articles dealing with Spain, with a particular strong rise starting in the early 1990s. Since this distribution is highly unequal, we have to normalise measures, such as frequency counts, and utilize text mining methods, such as STM, that can statistically accommodate unequal distributions.





Each article touches upon a range of different topics and frequently mentions other countries, i.e. Portugal, Greece, Italy or Ireland. In order to capture and analyse this country grouping heuristic, we create an additional variable called *country* that records if another PIGS country was mentioned in addition to Spain, which by definition, is always mentioned. Since Ireland was mentioned only very infrequently in comparison, references to this country were ultimately dropped,<sup>22</sup> meaning that the variable records either 'NA' (when only Spain is mentioned) or 'Greece,' 'Portugal,' or 'Italy' (when either of these countries was mentioned *in addition* to Spain). The year when each article was written is also recorded. Thus, for each *Zeit* article, there is metadata on the year it was written in and on associated PIGS countries.<sup>23</sup>

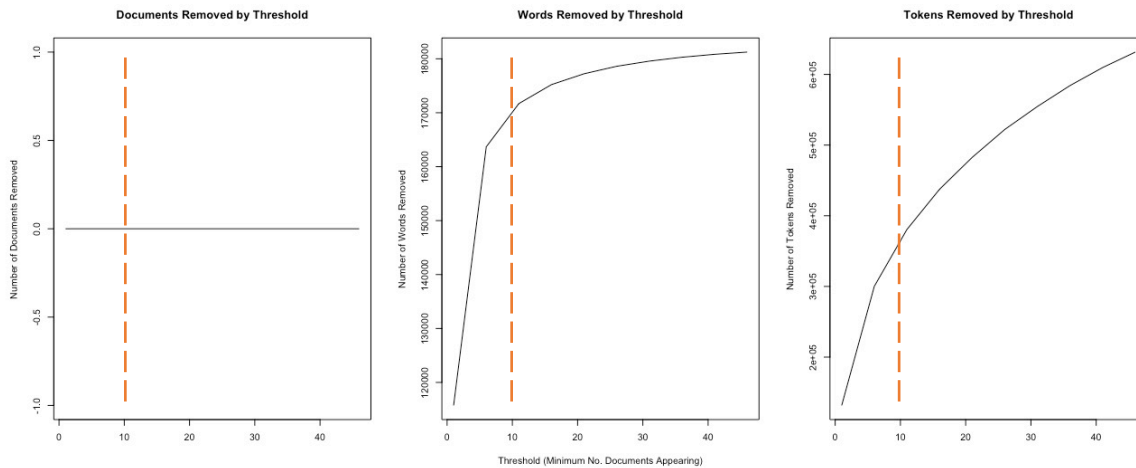
Within the field of text mining analysis, it is common to conduct some pre-processing of the semantic data before modelling it. The most common steps are stemming (reducing words to their root form), dropping punctuation and stop-word removal (common words that are not important for capturing a text's content, such as *the*, *is*, *at*). In our case, we convert all words of the *Zeit* articles to lower case, remove punctuation, remove stop-words, and remove numbers. Given that stemming algorithms available for the German language are far from perfect and often create more confusion than enhancing

<sup>22</sup> This is also the reason why we refer to the acronym PIGS and not PIIGS throughout the text.

<sup>23</sup> In our database, each article is a row in a .csv file, with the text contained in a variable called *text*.

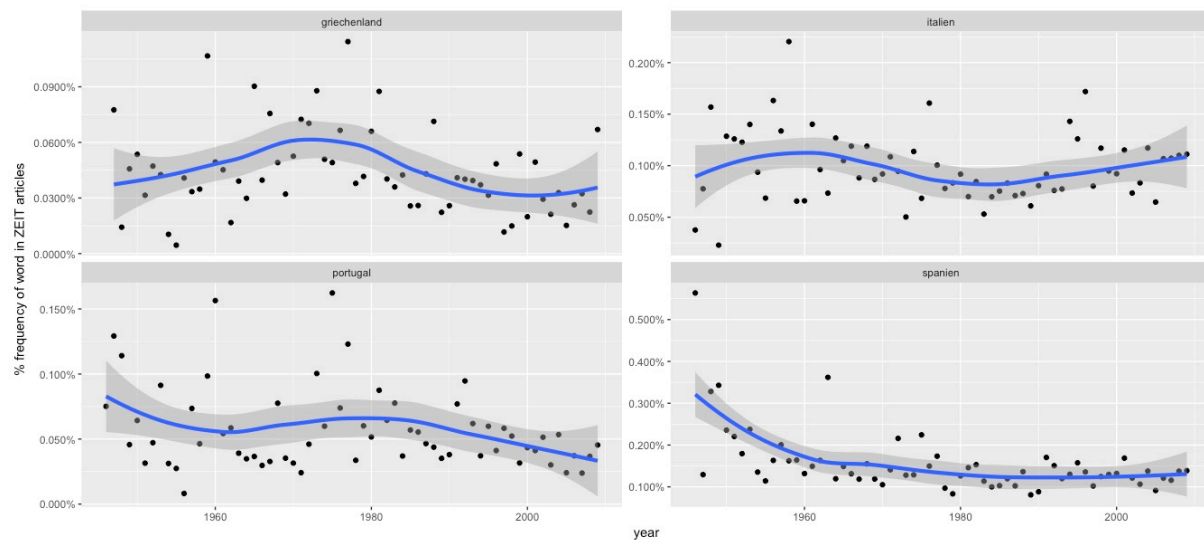
clarity, we do not perform stemming. We also do not define manual stop-words, in order to leave the corpus as authentic as possible.

After reading the data, researchers usually remove infrequent terms depending on a manually set parameter threshold, which decreases the computational time by several orders and ensures that the results are not affected by outliers. This lower threshold corresponds to the minimum number of *Zeit* articles a word needs to appear in, in order for the word to be kept within the vocabulary. The utility function, given in the figure below, plots the number of words and *Zeit* articles removed for different thresholds. Evaluating via the figure how many words would be removed from the dataset at each word threshold suggests that 10 articles is an effective threshold (see the orange dotted line). Selecting this threshold of 10 means that 170,636 of 185,157 terms (317,716 of 1,124,788 tokens) are removed due to lack of frequency. The final corpus of *Zeit* articles therefore consists of 2,443 articles, 14,521 terms and 807,072 tokens.



After reading in and processing the text data in the way described above, it is important to inspect features of the corpus and the associated vocabulary list to make sure that the documents have been correctly pre-processed. To this end, we present several descriptive statistics in the following paragraphs that give a good impression of the underlying semantic data and, at the same time, allow for a first analysis of the PIGS country group heuristic.

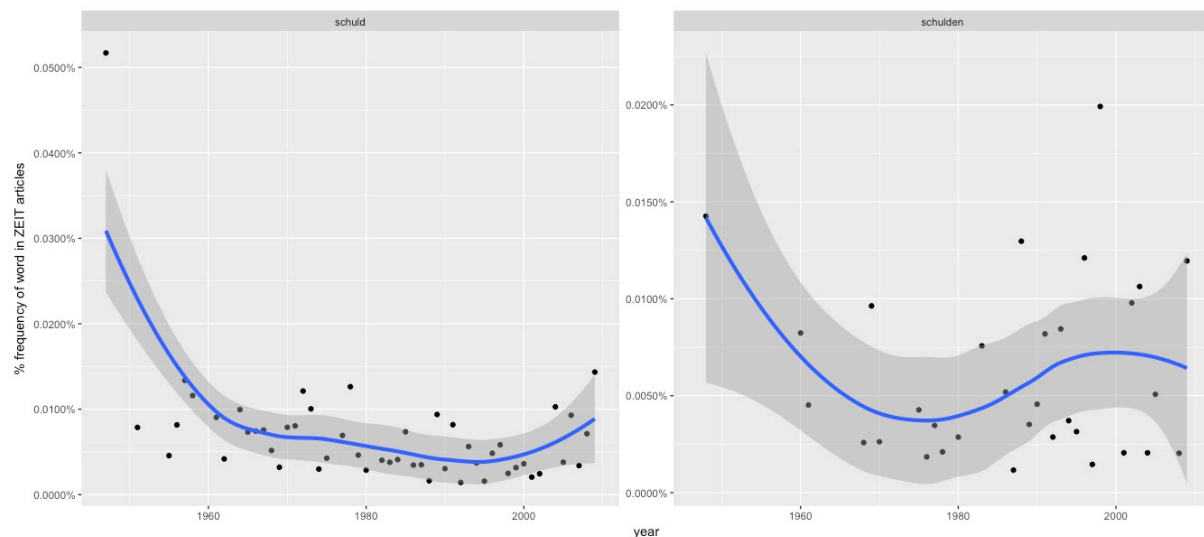
We start by plotting the relative frequency with which Spain and the PIG countries appear within the *Zeit* corpus (as a percentage of the total number of words used in the respective year). As can be seen from the figure below, Spain is initially particularly associated with its geographical neighbour Portugal. Besides the geographical proximity, this association in the early period was probably driven by the similarity of the political climate in these two countries at the time (both featured dictatorial systems, with Salazar in Portugal and Franco in Spain). From the 1970s onwards, Spain is increasingly associated with Greece. Again, this could be driven by political developments, namely the military coup in Greece that took place between 1967 and 1974. Over the whole period up until today, however, the most constant and strong association of Spain with another country was with Italy, the other large Southern economy in Europe. Given certain prominent discourses in the *Zeit* articles discussed later, such as football and European integration, this grouping of Spain and Italy is not surprising.<sup>24</sup>



The cultural divide between Northern and Southern EU members that was mentioned in the introduction became particularly visible in the crisis years due to Germany's opposition to fiscal transfers across

<sup>24</sup> As can be also seen from the figure, the frequency with which Spain was mentioned in articles over Spain declined over time.

Eurozone countries. This opposition has been explained with Germans' prudent attitude towards debt<sup>25</sup> and the important role of the 'Swabian housewife' metaphor in public discourse.<sup>26</sup> Thus, another semantic association with Spain, and Southern European economies in general, that might be explored in our corpus via frequency analysis is the German word for debt, *Schulden*. *Schuld* also means 'blame' or 'guilt', thus establishing a moral connotation. As can be seen in the figure below, we observe that this word appears in earlier decades, echoing Germany's special approach towards debt caused by the financial trauma of the interwar period. In the wake of massive reparation payments after WWI, the German currency had experienced a hyperinflation that survived in the political memory of the country.<sup>27</sup> Germany only agreed to eventually give up the *Mark* in favour of the Euro because the European Central Bank was designed, like the Bundesbank, as an anti-inflation institution.



<sup>25</sup> Hayo, B. and Neumeier, F., 'The social context for German economists: public attitudes towards macroeconomic policy in Germany', in G. Bratsiotis and D. Cobham (eds), *German Macro: How It's Different and Why That Matters* (Policy File, 2016), pp. 64-72.

<sup>26</sup> Anonymous, 'Hail, the Swabian housewife: The German mentality', *The Economist*, 410/8872 (2014), pp. 45-46.

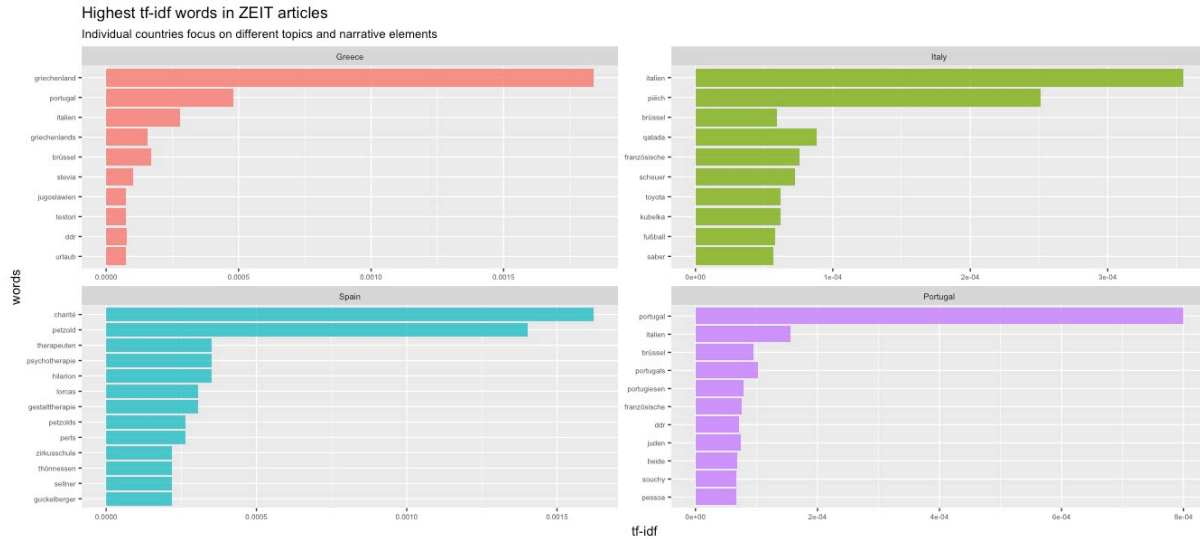
<sup>27</sup> For the role of Germany's central bank in the popularisation of this narrative, see: Mee, S., 'Monetary mythology: the West German central bank and historical narratives, 1948-78' (unpublished D.Phil. thesis, University of Oxford, 2016).

However, we also see from the figure that references to ‘Schulden’ are particularly numerous in the last few years of our period of investigation, that is in the course of European monetary integration and, ultimately, during the Eurozone crisis. This supports the notion, explained in more detail below, that the initial conception of Southern European countries as a homogenous, single group was driven by cultural factors (namely shared experiences of German tourists travelling southwards), while their latest incarnation, as the infamous PIGS group, was driven by an economic discourse on competitiveness and debt. Indeed, these economic discourses regarding the ECB and economic competitiveness feature prominently in the estimated STM discussed later, meaning they constitute a significant part of the *Zeit*’s reporting on Spain.

How important a word is in a *Zeit* article depends not only on its frequency in that article, but also how often it is commonly used. This is taken into account in the tf-idf method, with tf denoting ‘term frequency’ and idf denoting ‘inverse document frequency.’ When combining tf and idf, a term’s importance is adjusted for how rarely it is used. This means a text is analysed based on its position within the overall corpus: ‘The idea of tf-idf is to find the important words for the content of each document by decreasing the weight for commonly used words and increasing the weight for words that are not used very much in a collection or corpus of documents.’<sup>28</sup> A word in one of the *Zeit* articles is therefore important for this article if it appears frequently therein, but less frequently in other articles by way of comparison. The figure below plots the ten most important words, as measured by tf-idf, for each PIGS country.

---

<sup>28</sup> Silge, J. and Robinson, D., *Text Mining with R: A Tidy Approach* (eBook, 2017), <https://www.tidytextmining.com> (accessed: 06.09.2019).



Interestingly, the figure reveals that the three PIG countries are highly correlated with ‘Brüssel,’<sup>29</sup> pointing towards the role of the EU in these countries’ recent history and foreshadowing the difficulties that Southern European countries would experience during the Eurozone crisis when they repeatedly clashed with EU institutions over their debt management strategies. Moreover, we find evidence for the hypothesis that Germans’ perception of these countries was driven by ‘soft’ cultural factors such as car-enabled tourism or sport, given that the country profiles feature words related to these socio-cultural issues (‘urlaub’ in the case of Greece, ‘piech,’ ‘toyota’ and ‘fußball’ in the case of Italy). Finally, it should be noted that the rather curious word profile of Spain (prominently featuring medical terms such as ‘psychotherapie’ and ‘therapeuten’) is a statistical artefact, arising from the fact that our measures does not capture the most important words (as measured by tf-idf) for *all* articles dealing with Spain (since this would equal the whole corpus), but captures the most important words that distinguish the *Zeit* articles that feature *only* Spain in contrast with all other articles (that include PIG countries).

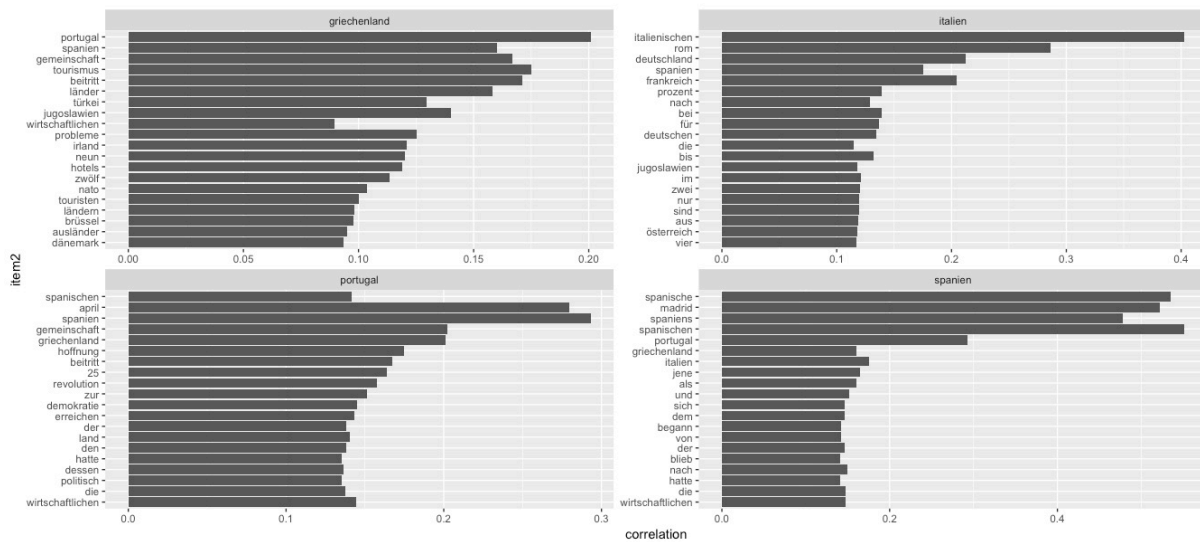
Finally, we can investigate the country group heuristic by analysing correlations between the PIGS countries’ names and other words used in the *Zeit* articles. While simply calculating the most common co-occurring words (such as ‘Spanien’ and ‘spanische’) is not particularly meaningful since they are often also the most common individual words, examining correlation among words indicates how often they

<sup>29</sup> Here, and in the remainder of this paper, we use single inverted commas to refer to explicit terms in the corpus (so-called token) in order to distinguish them from the estimated topics (which will be shown in *italics* and capitalised).

appear together relative to how often they appear separately. In particular, we investigate word correlations via the ‘phi coefficient,’ a common measure for binary correlation. The focus of the phi coefficient is how much more likely it is that either both word X and Y appear, or that neither do, than that one appears without the other.<sup>30</sup> The phi coefficient can be written as:

$$\phi = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{.1}n_{.0}n_{0.1}n_{1.0}}}$$

where  $n_{11}$  represents the number of *Zeit* articles where both word X and word Y appear,  $n_{00}$  the number where neither appears, and  $n_{10}$  and  $n_{01}$  the cases where one appears without the other, whereas  $n_{.1}$  denotes all articles that contain Y and  $n_{.0}$  all that do not contain word Y. We calculate the phi coefficient between words based on how often they appear in the same *Zeit* article. In particular, we pick the four PIGS countries and find the other words most associated with them (figure below).



From the figure, we see that Greece is associated with tourism and economic problems, Italy with Rome and Germany, Portugal with the European Union and revolution, and Spain with Madrid and economic aspects. Most interesting for our purposes is the fact that all PIGS countries have high correlations with each other: Greece is strongly correlated with Spain and Ireland (and also Turkey, Yugoslavia and Denmark), Italy with Spain (as well as Germany, France and Yugoslavia), Portugal with Spain and Greece,

<sup>30</sup> Silge, J. and Robinson, D., *Text Mining with R: A Tidy Approach* (eBook, 2017), <https://www.tidytextmining.com> (accessed: 06.09.2019).

and Spain with all three PIG countries. This implies that German media, as approximated through the *Zeit* corpus, often referred explicitly to other Southern European countries when discussing one of the PIGS countries, and particularly when discussing Spain, thereby perpetuating this country group heuristic. Having surveyed the information that can be obtained through standard descriptive statistics and correlation analysis, we are ready to estimate a STM.

### Structural Topic modeling

We begin by providing a brief overview of the STM model.<sup>31</sup> Like the initial topic models designed by Blei et al.,<sup>32</sup> STM is a generative model of word counts. In short, this means that the algorithm defines a data generating process for each document and then uses the semantic data, as captured in our *Zeit* articles, to find the most likely values for the parameters within the model. This generative process can be understood as a three-step procedure: each *Zeit* article is assumed to exhibit the topics in different proportion (step 1), and each word in each article is drawn from one of these topics (step 3), where the selected topic is chosen from the per-article distribution over topics (step 2).<sup>33</sup> Given the number of topics that should be produced, the model places together terms that appear in the same article more frequently than one would expect by chance. In an iterative process, each word is assigned to a topic, which is therefore defined as a mixture over words, with each word having a probability of belonging to a topic. Thus, a *Zeit* article is understood as comprising a mixture of topics, meaning that a single article can consist of several different topics to a varying degree.<sup>34</sup>

So far, this resembles the typical structure of a classic topic model. The additional advantage introduced by STM is the fact that it allows to incorporate metadata into the estimation framework. This can happen in two ways that can be classified as either affecting ‘topical prevalence’ or affecting ‘topical content.’

---

<sup>31</sup> For technical details, see: Roberts, M.E., Stewart, B.M. and Tingley, D., ‘stm: R Package for Structural Topic models’, *Journal of Statistical Software*, available online: <https://cran.r-project.org/web/packages/stm/vignettes/stmVignette.pdf> (accessed: 06.09.2019).

<sup>32</sup> Blei, D.M., Ng, A., Jordan, M., ‘Latent Dirichlet Allocation’, *Journal of Machine Learning Research*, 3 (2003), pp. 993-1022.

<sup>33</sup> Overview: Blei, D.M., ‘Topic modeling and Digital Humanities’, *Journal of Digital Humanities*, 2/1 (2012), <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/> (accessed: 06.09.2019).

<sup>34</sup> Mathematically, this means that the sum of the topic proportions across all topics for a document is one, and the sum of word probabilities for a given topic is one.



Metadata covariates for topical prevalence allow the observed metadata to influence the frequency with which a topic is discussed, whereas topical content covariates allow the observed metadata to influence the word rate use within a given topic (meaning how a particular topic is discussed).<sup>35</sup>

Since different *Zeit* articles are associated with different years, and thus different cultural, political, and economic contexts, it is natural to want to allow this prevalence to vary with the metadata that we have about our articles. We will let prevalence be a function of the *country* variable, which codes references to other PIG countries, and the variable *year*, which is an integer measure of years running from the first article written in 1946 to the last available articles published in 2009. We enter in the variables additively, by allowing for the *year* variable to have a non-linear relationship in the topic estimation stage.<sup>36</sup> Moreover, we allow the *country* variable to also affect topical content.

Before the algorithm can start its work, we have to specify the number of topics to be estimated. Which number is ideal? Recently, several metrics have been proposed with which to identify the correct number of topics.<sup>37</sup> However, they provide the optimal number of topics only in a statistical sense. Most researchers in the social sciences therefore run the algorithm several times and compare the respective outputs.<sup>38</sup> In short, there is not a ‘right’ answer to the number of topics that are appropriate for a given corpus and ultimately, one should decide for the most coherent and helpful output given the initial research question. Following several estimations with different topic numbers (5, 10, 15, 20, 50), we opt for a 15 topic STM model. However, it is noteworthy that the general conclusions are not affected by selecting this specific model.<sup>39</sup>

---

<sup>35</sup> Note that STM allows using topical prevalence covariates, a topical content covariate, both, or neither.

<sup>36</sup> Technically speaking, the *stm* package includes a convenience function, which selects a flexible b-spline basis. In our model, we allow for the variable *year* to be estimated with a spline.

<sup>37</sup> See, e.g., the R package *ldatuning* developed by Murzintcev Nikita: <https://cran.r-project.org/web/packages/ldatuning/vignettes/topics.html> (accessed: 06.09.2019).

<sup>38</sup> Wehrheim, L., ‘Economic history goes digital: topic modeling the Journal of Economic History’, *Cliometrica* 13 (2019), p. 113; also: Ferri, P., Lusiani, M. and Pareschi, L., ‘Accounting for Accounting History: A Topic modeling Approach (1996–2015)’, *Accounting History*, 23/1-2 (2018), pp. 173-205.

<sup>39</sup> A smaller or larger number of topics merely increases the number of vague topics with conflicting words that are difficult to interpret (for a discussion of these vague topics, see below in the main text). However, all the estimated models had a set of certain common topics dealing with cultural, political and economic factors of Spain’s image and this set of discourses constituted the most significant proportion of the *Zeit* corpus.

The model is set to run for a maximum of 75 iterations using a self-selected seed. Convergence is monitored by the change in the approximate variational lower bound. Once the bound has a small enough change between iterations, the model is considered converged. Most topic models, including STMs, are sensitive to initialization, meaning that the results obtained from the estimation could vary in line with the starting values of the parameters (e.g. the distribution over words for a particular topic).<sup>40</sup> To deal with this, we use so-called ‘spectral initialization’ based on the method of moments,<sup>41</sup> which guarantees that irrespective of the seed that is set, the same output will be generated. We argue that this initialization method should be also used by other researchers who would like to implement topic modelling in their research since the replicability of the results facilitates the sharing of data as well as the comparison of results and, overall, guarantees a more transparent academic discourse.

## RESULTS I: TOPICS

We explore the topics that have been estimated via two approaches. First, we aim to understand the content of the fifteen topics, and then we visualise their respective proportion within the corpus of *Zeit* articles. As stated earlier, topic models treat topics as distributions over words. Accordingly, the computational output consists of the highest probability words associated with each topic. However, these groups lack any kind of label. Historians usually assign a single label to each topic in order to facilitate an aggregate analysis of the corpus’ content. For instance, in mining newspaper articles of the *Richmond Daily Dispatch*, Nelson looks at a topic with words like ‘treasury,’ ‘bonds,’ and ‘interest,’ and identifies it with the label *War bonds*.<sup>42</sup> Likewise, by reading through the estimated word lists as well as examining actual articles that are estimated to be highly associated with each topic, we are able to summarise each *Zeit* topic with a manually specified label. The results are given in the table below.

---

<sup>40</sup> This is due to the fact that as mixed-membership models, their posterior is intractable and non-convex, which creates a multimodal estimation problem.

<sup>41</sup> Arora, S., Ge, R., Halpern, Y., Mimno, D., Moitra, A., Sontag, D., Wu, Y. and Zhu, M., ‘A Practical Algorithm for Topic Modeling with Provable Guarantees’, in S. Dasgupta and D. McAllester (eds.), *Proceedings of the 30th International Conference on Machine Learning*, 28 (2013), pp. 280-288.

<sup>42</sup> Nelson, R.K., ‘Mining the Dispatch’, *Blog*: <http://dsl.richmond.edu/dispatch/Topics/view/28> (accessed: 16.11.2017).

UPIER WORKING PAPER Vol. 16

A topic model with 15 topics, 2,443 documents and a 14,521-word dictionary		
Topic	Topic words (high probability)	Label
1	buchmesse, weine, übersetzungen, flüssigkeit, flaschen, jahrgang, kompetenzen	Culture I (books and wine)
2	erklärung, regierungen, angeblich, empört	Rhetoric terms
3	ezb, geldpolitik, ews, währungsunion, euro-währung, trichet, konvergenzkriterien	European monetary policy
4	viertelfinale, mönchengladbach, spielern, stürmer, borussia, brasilianer, dfb	Football
5	gedichten, heiratete, poetischen, romane, verse, michelangelo, kloster	Culture II (literature and art)
6	arbeitslosenzahl, rückgang, anleger, arbeitskosten, sozialprodukts, jahresbeginn, unangefochten	Economic competitiveness
7	putsch, fidel, islamisten, hussein, anarchisten, militanten, cardinal	Political turmoil
8	fluggäste, ltu, lufthansa, verkehrsminister, airways, condor, autofahrer	Travels
9	philosophische, großfamilie, training, wahrnehmen, jakob, dramen, lyric	Diverse terms I
10	steinbrück, angela, kohlendioxid, potenzial, ökonom, passt, internet	German politics
11	vollversammlung, truman, eisenhower, blei, uno, imperialismus, großmächten	International Relations
12	bungalows, trip, kreuzfahrt, halbpension, postfach, kreuzfahrten, hummel	Tourism
13	eg-kommission, eurokraten, schutzmaßnahmen, eg-präsidentschaft, bangemann, eg-kommissar, thatcher	European integration

14	chruschtschows, pluralismus, politbüro, imperialismus, großmächten, supermächte, chruschtschow	Cold War
15	vorhin, bruttosozialprodukts, dahrendorf, labour, technologische, unsere, schmidt	Diverse terms II

Surveying the topic labels given in the table, we can identify three discourses within the *Zeit* newspaper corpus, spanning 2,443 articles and almost seven decades. Most importantly for our purpose, we can identify a large *cultural discourse* encompassing topics mostly related to Germans travelling abroad. To begin with, the *Tourism* topic (topic 12) encompasses revealing terms such as ‘trip,’ ‘kreuzfahrten’ and ‘halbpension.’ More narrowly, the related topic *Travel* (topic 8) focuses on the act of transportation, most commonly via plane (‘fluggäste,’ ‘Lufthansa,’ ‘airways’), but also via car (‘autofahrer’). The two topics labelled *Culture I* and *II* (topics 1 and 5) deal with culture in the narrow sense, that is art (‘buchmesse,’ ‘gedichte,’ ‘romane,’ ‘michelangelo’) and food (‘weine,’ ‘jahrgang’). As might be expected, the German view of Southern Europe, and particularly Spain and Italy, was also shaped by these countries’ great *Football* (topic 4) teams, which often clashed with the German national team (‘dfb’) or national clubs (‘borussia,’ ‘mönchengladbach’) in the different competitions (‘viertelfinale’), thereby creating myths and cultural heuristics that were passed on to subsequent generations of sport enthusiasts.

A closer look into the corpus shows that the estimated *Culture* topics accurately capture the content of many articles.<sup>43</sup> The economic upswing, which began with the establishment of the social market economy by Ludwig Erhard and the incipient economic miracle, led to a tourism euphoria in Germany, which found its clear expression in many *Zeit* articles. ‘Everyone is hungry for sun; this summer more than ever!’ the director of a Hamburg travel agency is quoted in a 1962 article, explaining the fact that special trains and ships, charter planes and coaches to the South were occupied to the last place, and that of approximately 320,000 holiday travellers in that year almost two thirds went abroad, especially to Italy

---

<sup>43</sup> The methodology that was followed for this and the subsequent close reading exercises was as following: As a starting point, we read all articles that contained the most significant words of the respective topics as estimated by the STM and after analysing these articles, we had gained a sufficient understanding of the respective topics allowing us to carry on the corpus search with additional, manually created, keywords.

and Spain.<sup>44</sup> A good example of how this euphoric mood was linked to the PIGS countries is the following quote from a 1960 article, which announces ‘even in December we do not have to turn our backs on Europe if we want to romp around in the water,’ followed by a survey of potential destinations that prominently put together the PIGS countries’ tourism destinations:

‘In the Christmas month, the coasts of Greece were waiting for the water rides from the north. At the gulfs of Arkadia, Koroni and Nauplia (Peloponnese) there are still many sunny beaches where we can welcome Santa Claus in a bathing costume. Also, the islands off the coast of the Peloponnese, like Kephallinia, Zakynthos and Kythira, promise Christmassy bathing pleasures. The Gulf of Taranto and Catanzaro at the tip of Italy's boots as well as the southern coasts of Sicily also offer pleasantly tempered bathing water. We can also celebrate New Year's Eve on Spain's south coast west of Almeria and on Portugal's south coast between the Spanish border and Faro if the weather is favourable.’<sup>45</sup>

Similar tourism-related articles can be found in many other subsequent years as well, particularly during the 1960s and early 1970s (in line with the estimated temporal pattern shown later), often written by the journalist Horst Hachmann.<sup>46</sup>

---

<sup>44</sup> “‘Alles hungert nach Sonne; in diesem Sommer mehr denn je!’ So erklärt der Direktor eines Hamburger Reisebüros die Tatsache, daß Sonderzüge und Schiffe, Charterflugzeuge und Reisebusse in den Süden bis auf den letzten Platz belegt sind, und daß von ungefähr 320 000 Ferienreisenden in diesem Jahr fast zwei Drittel ins Ausland und hauptsächlich nach Österreich, Italien und Spanien führen.’ Schnabel, C., ‘Die Reisebüros sind jetzt leer...’, *Die Zeit* (1962). As primary sources, all *Zeit* articles quoted from our corpus are given in this abbreviated version and are not included in the bibliography.

<sup>45</sup> ‘Selbst im Dezember brauchen wir Europa noch nicht den Rücken zu kehren, wenn wir uns im Wasser tummeln wollen. Im Weihnachtsmonat warten die Küsten Griechenlands auf die Wasserritten aus dem Norden. An den Golfen von Arkadia, Koroni und Nauplia (Peloponnes) gibt es noch viele sonnenwarme Strände, wo wir den Weihnachtsmann im Badekostüm begrüßen können. Auch die dem Peloponnes vorgelagerten Inseln, wie Kephallinia, Zakynthos und Kythira, verheißen weihnachtliche Badefreuden. Ebenso warten der Golf von Tarent und Catanzaro an der Stiefelspitze Italiens sowie die Südküsten Siziliens mit angenehm temperiertem Badewasser auf. Silvester können wir im Badedreß auch noch an Spaniens Südküste westlich von Almeria feiern und auch an Portugals Südküste zwischen der spanischen Grenze und Faro, wenn das Wetter günstig ist.’ Westphal, P., ‘Europas südliche Meere laden zum Bade’, *Die Zeit* (1960).

<sup>46</sup> For instance: ehö, ‘Auch die stets lächelnden Optimisten unter den Reisebürokaufleuten...’, *Die Zeit* (1964); Ranft, F., ‘Offensichtlich besteht bei vielen jungen Leuten...’, *Die Zeit* (1964); Anonymous, ‘Texas Ranch und Ol’ Man River’, *Die Zeit* (1965); Baade, F., ‘Zehn Monate Badesaison – unbekannte Riviera Antike in Superlativen’, *Die Zeit* (1965); Anonymous, ‘Kindermenüs in der Luft’, *Die Zeit* (1966); Hachmann, H., ‘Drei Merkmale...’, *Die Zeit* (1966); Hachmann, H., ‘Der reisende Michel...’, *Die Zeit* (1966); Hachmann, H., ‘Sollte die touristische Hochrechnung für das Jahr 1967 aufgehen...’, *Die Zeit* (1967); Anonymous, ‘Offenbar haben Jugendreisen hierzulande kein besonders gutes Image...’, *Die Zeit* (1970); Sachse, U., ‘Der Optimismus der Reiseveranstalter...’, *Die Zeit* (1971); bo, ‘Wer je mit Schiffen die griechische Inselwelt bereiste...’, *Die Zeit* (1973).

In the literature on tourism, it is well acknowledged that certain nationalities have preferences for individual places, particularly amongst migrants of a specific age.<sup>47</sup> Historically, besides the Costa Blanca around Alicante, the Balearic Islands and the Canary Islands, the Costa del Sol's coastal areas became the most popular Spanish destinations for German tourists. A good example is the site Torrox on the Costa del Sol that was opened up for tourism in the 1970s when, on the initiative of a German construction company, blocks of flats were built that were sold exclusively to Germans. Since then, the town has offered generations of German tourists, and older German migrants, the opportunity to pursue a leisure-oriented lifestyle and has therefore developed into a popular destination for German tourists.<sup>48</sup> As the above quotation from the 1960 *Zeit* article illustrates, these touristic experiences were well reflected in the German media, which initially grouped these countries together due to their geographical suitability for sunny holidays.

Of particular interest is the fact that some of these tourism-centred *Zeit* articles disprove the common image of the young German holidaymaker who, in the upcoming age of mass tourism, only aimed to achieve a tan as symbol of the 'successful' holiday, and instead point towards cultural exchanges that might have enriched the German travellers' perception of Southern Europe. As the writer and *Zeit* journalist Ferdinand Ranft noted in 1966: 'A not inconsiderable part of the young participants in organised holidays wants activities of a sporting, cultural and foreign nature,' and in his article, he lists several examples of how German travellers could realise these experiences.<sup>49</sup> This perception of the German tourist travelling abroad in order to enrich his cultural image of other countries still featured in a 1971 article that described study trips to Spain, Italy, and Greece: 'The trips are not lazy fun, but exhausting educational trips. Travellers not only want to see the world, they also want to understand it.'<sup>50</sup> During these trips, Germans compared the cultural images of Southern Europe that they had acquired during their childhood through hearsay or literature with the contemporary realities, as described in this 1972 article: 'All the cities on the Mediterranean are the scenes of images and books

---

<sup>47</sup> Breuer, T, 'Ein Dauerplatz an der Sonne', *Praxis Geographie*, 3 (2002), p. 22.

<sup>48</sup> Kordel, S., 'Lifestyle-Mobilitäten deutscher Senioren in Spanien. Das Beispiel der Gemeinde Torrox an der Costa del Sol', *Mitteilungen der Fränkischen Geographischen Gesellschaft*, 58 (2011), pp. 53-66.

<sup>49</sup> Ranft, F., 'Das gängige Bild des jungen Urlaubers', *Die Zeit* (1966).

<sup>50</sup> 'Die Reisen sind kein Faulenzervergnügen, sondern anstrengende Bildungsreisen. Die Reisenden wollen die Welt nicht nur sehen, sondern auch verstehen.' bo, 'Die Reisen sind kein Faulenzervergnügen...', *Die Zeit* (1971).

that we, as children, have seen and – more and more demandingly – devoured with our torches in bed. There are accordingly many conceptions.<sup>51</sup> While this confrontation with reality could sometimes lead to disappointment, as the continuation of the quote illustrates, it definitely confirms a mental process of forming country group heuristics based on cultural features:

‘Greece and its ports are now linked to the image of the generals. Still for a somewhat older generation, it was the glory of Hellas associated with Winkelmann, Goethe and neo-Humanism. Or Spain: Today Torremolinos, Marbella and the complaint of Peter Neckermann, who expresses concern about the construction boom in the Mediterranean region, where without a vote concrete silos and hotel capacities are created. [...] For an older generation, the image of Spain was still marked by the Spanish Civil War and Hemingway’s books, the Andalusian mountain nest Ronda, Pamplona and Barcelona were under discussion.’<sup>52</sup>

Besides this cultural discourse, the topics estimated on the basis of our *Zeit* corpus reflect the most important geo-political developments during the period of investigation, thus forming a large *political discourse*. Starting with the national level, several themes debated within *German Politics* (topic 10) were featured in the *Zeit* articles, explicitly naming German politicians (‘angela,’ ‘steinbrück’). *Political Turmoil* (topic 7) was associated with developments in more distant countries (‘fidel,’ ‘husein’) and a diverse set of extreme political goals (‘islamisten,’ ‘anarchisten,’ ‘militanten’), which nevertheless bear some relevance for the PIGS heuristic given that Eurocommunism influenced especially the communist parties of Spain and Italy. The *Cold War* (topic 14) centring around the Soviet-Union (‘chruschtschow,’ ‘politbüro’) and the US (‘Truman,’ ‘Eisenhower’ in topic 11) was, of course, an important topic for several decades and also shaped the way that *International Relations* (topic 11) were framed (‘imperialismus,’ ‘großmächten’). Towards the end of our period, the development of European political institutions (‘eg-

---

<sup>51</sup> ‘Alle Städte am Mittelmeer sind Schauplätze der Bilder und Bücher, die wir, als Kind, schon mit der Taschenlampe im Bett, und dann, immer anspruchsvoller, gesehen und verschlungen haben. Es gibt entsprechend viele Vorstellungen.’ Anonymous, ‘Alle Städte am Mittelmeer sind Schauplätze...’, *Die Zeit* (1972).

<sup>52</sup> ‘Griechenland samt seinen Häfen ist heute mit dem Image der Generale verbunden. Noch für eine etwas ältere Generation war es die Winkelmannsche, die Goethesche und neuhumanistische Glorie von Hellas. Oder Spanien: Heute Torremolinos, Marbella und die Klage von Peter Neckermann, der Sorge äußert hinsichtlich des Bauboomes im Mittelmeerraum, wo ohne Abstimmung Betonsilos und Hotelkapazitäten entstehen. [...] Für eine ältere Generation war das Bild von Spanien noch durch den spanischen Bürgerkrieg und die Bücher Hemingways geprägt, waren das andalusische Bergnest Ronda, Pamplona und Barcelona im Gespräch.’ Anonymous, ‘Alle Städte am Mittelmeer sind Schauplätze...’, *Die Zeit* (1972).

kommission’, ‘eg-kommissar’, ‘eg-präsidentschaft’) that were related to the process of *European Integration* (topic 13) became increasingly relevant, but was also received partly critically as being too technocratic and elite-centred (‘eurokraten’).

Anticipating the results from plotting these topics’ proportion over time, we focus on the two most relevant topics in the *Zeit*’s political discourse, namely *International Relations* (topic 11) and *European Integration* (topic 13). Again, reading through the original *Zeit* articles confirm that the choices of the STM do indeed represent the most important discourses with regard to political issues. However, the first topic’s link to the PIGS countries seems to be rather coincidental: many articles deal with the danger for peace posed by the Soviet Union and use a comparative methodology to analyse the different influence of communist groups in European countries. For instance, a 1968 article argues that this influence is relatively weak because ‘in the industrialized European countries, the communists play very different roles: in the government (Finland), as the main political forces (France and Italy), as insignificant political sects (Denmark and New Zealand), or in illegality (Spain, Portugal, Greece, Germany).’<sup>53</sup> Three years later, an article described the presence of ‘three major communist parties in Western Europe (in Italy, France and Spain),’<sup>54</sup> thereby again forming a country group heuristic simply by focusing on the role of the communists in Southern European countries and identifying a comparable pattern in this region.

More revealing are the articles dealing with *European Integration* (topic 13), surprisingly they anticipate many of the arguments that were later advanced in the German public discourse on PIGS countries during the Eurozone crisis. Already at the early stage of the European Community, an article in 1975 by Rolf Zundel explains to the readers of the *Zeit* that ‘European policy is by no means just about whether

---

<sup>53</sup> ‘In den Industrieländern spielen die Kommunisten ganz verschiedene Rollen: in der Regierung (Finnland), als maßgebliche politische Kräfte (Frankreich und Italien), als unbedeutende politische Sekten (Dänemark und Neuseeland) oder aber in der Illegalität (Spanien, Portugal, Griechenland, Bundesrepublik).’ Anonymous, ‘Gegenwärtig gibt es in der Welt 94 kommunistische Parteien...’, *Die Zeit* (1968).

<sup>54</sup> ‘Wenn man berücksichtigt, daß die Sozialdemokraten auch in den skandinavischen Ländern einen kolossalen Einfluß haben und daß sie in Schweden sogar im vierten Jahrzehnt an der Macht sind, wenn man ferner berücksichtigt, daß in Westeuropa drei große kommunistische Parteien existieren (in Italien, Frankreich und Spanien), so muß man folgern, daß es zur Zeit auf dem europäischen Kontinent genügend Kräfte gibt, die sowohl einen Kalten Krieg als auch einen richtigen Krieg verhüten könnten.’ Anonymous, ‘Einige sowjetische Politiker und Strategen...’, *Die Zeit* (1971).



Bonn finds the right tone for its partners,' but primarily about a 'methodological dispute' evolving around two problems:

'On the one hand, it does not always benefit the recipient countries if German money is used to plug holes in the budget. In this way, it is feared in Bonn, the inefficiency will be made permanent. On the other hand, the Ministry of Finance and the Chancellery are forced more than ever to think about the limits of German economic performance. Both considerations are working in the same direction: the dosage and conditioning of Bonn payments is being considered more closely.'<sup>55</sup>

This resembles the so-called moral hazard argument put forward by German economists and politicians during the Eurozone crisis, who worried that avoiding bailouts would set a bad example and encourage poor behaviour among other actors.<sup>56</sup> Back in the 1970s, *Zeit* authors also understood that the heterogenous nature of the different European economies made them very different from the standard model of an optimal currency area:<sup>57</sup>

'The objection in principle to the continued enlargement of the Community focuses on the steep gap between the powerful and the underdeveloped members. The sluggishness of the decision-making process is a cause for concern, as the compromise would become more and more expensive and the substance increasingly lean.'<sup>58</sup>

---

<sup>55</sup> 'Zum einen nützt es den Empfängerländern nicht immer, wenn deutsches Geld zum Löcherstopfen im Budget verwendet wird. Auf diese Weise, so befürchtet man in Bonn, werde der Schlendrian dauerhaft gemacht. Zum anderen ist man im Finanzministerium und im Kanzleramt mehr denn je gezwungen, an die Grenzen deutscher Leistungsfähigkeit zu denken. Beide Überlegungen wirken in die gleiche Richtung: die Dosierung und Konditionierung der Bonner Zahlungen wird genauer bedacht'. Zundel, R., 'Bonns neue Rolle in der Außenpolitik', *Die Zeit* (1975).

<sup>56</sup> James, H., 'Rule Germania', in T. Beck and H.-H. Kotz (eds), *Ordoliberalism: A German oddity?* (London, 2017), pp. 26ff.

<sup>57</sup> The classic reference is, of course, Robert Mundell: Mundell, R., 'Currency Areas, Common Currencies and EMU', *The American Economic Review*, 87/2 (1997), pp. 214-216.

<sup>58</sup> 'Der prinzipielle Einwand gegen eine fortdauernde Erweiterung der Gemeinschaft konzentriert sich auf das steile Gefälle zwischen leistungsstarken und unterentwickelten: Mitgliedern. Sorgen bereitet die Schwerfälligkeit der Beschlußfassung, weil der Kompromiß immer teurer und in der Substanz immer magerer ausfallen würde'. Becker, K., 'Portugals EG-Kandidatur erzwingt eine historische Entscheidung', *Die Zeit* (1977).

It is exactly this economic divergence that is seen nowadays as the origin of the Eurozone's problems, namely by leading to social differences and to discrepancies in political interests regarding the further direction of economic integration.<sup>59</sup> The Germans' later distaste for so-called Eurobonds is likewise echoed in the *Zeit* articles paralleling the course of European integration. When discussing the potential membership of Portugal, the *Zeit* author concludes: 'The transition to majority decision-making is pure theory. Nobody wants to be outnumbered, not even Bonn; because then, to cite a particularly deterrent example, the euro would threaten the socialization of our foreign exchange stocks.'<sup>60</sup>

Lastly, the estimated topics allow us to identify an *economic discourse* in the *Zeit* articles. At first sight, this discourse seems surprisingly slight, at least in comparison with the larger number of topics concerning cultural and political issues, but as we have seen, many political articles do touch upon economic issues. More specifically, the economic discourse encompasses two topics, namely *Economic Competitiveness* (topic 3), consisting of words such as 'arbeitskosten' and 'sozialprodukt,' and *European Monetary Policy* (topic 6), defined for instance by 'geldpolitik' and 'währungsunion.' The latter topic is strongly characterised by its European Community dimension ('ezb') and the convergence criteria associated with membership in this community ('konvergenzkriterien'). The wish to promote convergence has always played a central role in the historical development of European monetary integration.<sup>61</sup> In the Delors Report (1989) we find an early discussion of the need to promote greater convergence in economic performance and living standards, as well as in economic policies, in order to clear the path towards a monetary union.<sup>62</sup> However, with the creation of the Eurozone, the PIGS countries lost a crucial option for steering economic policy, namely the option of periodically devaluing their own currency; this would make it much more difficult to achieve this convergence in the long run.<sup>63</sup>

---

<sup>59</sup> Tokarski, P., 'Divergence and Diversity in the Euro Area. The Case of Germany, France and Italy', *SWP Research Paper*, 06 (2019), <https://www.swp-berlin.org/10.18449/2019RP06/> (accessed: 06.09.2019).

<sup>60</sup> 'Der Übergang zur Mehrheitsentscheidung ist reine Theorie. Niemand will sich majorisieren lassen, auch Bonn nicht; denn dann drohte, um ein besonders abschreckendes Beispiel zu nennen, die Euro-Sozialisierung unserer Devisenbestände.' Becker, K., 'Portugals EG-Kandidatur erzwingt eine historische Entscheidung', *Die Zeit* (1977).

<sup>61</sup> Franks, J., Barkbu, B., Blavy, R., Oman, W. and Schoelermann, H., 'Economic Convergence in the Euro Area: Coming Together or Drifting Apart?', *IMF Working Paper* (January 2018), <https://www.imf.org/en/Publications/WP/Issues/2018/01/23/Economic-Convergence-in-the-Euro-Area-Coming-Together-or-Drifting-Apart-45575> (accessed: 06.09.2019).

<sup>62</sup> European Council and Committee for the Study of Economic and Monetary Union, *Report on economic and monetary union in the European Community* (Unipub, 1989).

<sup>63</sup> Tokarski, P., 'Divergence and Diversity in the Euro Area. The Case of Germany, France and Italy', *SWP Research Paper*, 06 (2019), <https://www.swp-berlin.org/10.18449/2019RP06/> (accessed: 06.09.2019).

This economic dilemma is also discussed in the *Zeit* articles, as another close reading exercise confirms. The context is given by the 1992 Maastricht Treaty that established convergence criteria for countries to join the euro, by focusing on nominal and fiscal indicators of harmonization, including inflation, long-term interest rates, exchange rate stability, the fiscal deficit, and the government debt-to-GDP ratio. Between 1991 and 1998, several *Zeit* articles described and analysed these Euro convergence criteria.<sup>64</sup> In this reporting, the authors generally applauded the criteria, which they interpreted as ‘mainly strict requirements in terms of price and exchange rate stability and budgetary discipline,’ concluding that the ‘Federal Government and the Bundesbank were largely able to assert themselves with their high demands on the soundness of their partners.’<sup>65</sup> From the start, however, *Zeit* authors were sceptical of the PIGS countries’ readiness for fulfilling these criteria and joining the Euro-area: ‘Even in the longer run, some countries already seem to have no chance. The most striking examples: Belgium, Greece and Italy have such high debts that it will be impossible for them to reduce them to the required level in a few years’ time. [...] Portugal, Ireland and Spain have little chance. They would have to make huge savings and thus risk enormous social tensions.’<sup>66</sup> In short, the transformation from culturally acclaimed tourist destination to a negative country group heuristic based on economic attributes was achieved during the 1990s. The rise of this new image of the PIGS would ultimately lead to pessimistic reporting on the ‘spectre of monetary union.’<sup>67</sup>

---

<sup>64</sup> E.g. Anonymous, ‘Europas Regierungschefs...’, *Die Zeit* (1991); Anonymous, ‘Die Richter in den roten Roben haben gesprochen...’, *Die Zeit* (1993); Cartellieri, U., ‘Trotz Verfassungsgerichtsurteil, trotz politischer Willensbekundungen...’, *Die Zeit* (1993); Marsh, D., ‘Entrüstung in Rom, Madrid und Kopenhagen...’, *Die Zeit* (1994); Anonymous, ‘Die Turbulenzen des neuen Jahres...’, *Die Zeit* (1995); Anonymous, ‘Schreckgespenst Währungsunion’, *Die Zeit* (1995); Anonymous, ‘Sind Europa, die Europäische Währungsunion und die europäische Einheitswährung nur noch Themen für Märchenerzähler?...’, *Die Zeit* (1996); Anonymous, ‘Anfang Mai werden die Staats- und Regierungschefs der EU...’, *Die Zeit* (1998).

<sup>65</sup> ‘Dabei handelt es sich vor allem um strenge Anforderungen an Preis- und Wechselkursstabilität und an die Haushaltsdisziplin. [...] Bundesregierung und Bundesbank konnten sich mit ihren hohen Anforderungen an die Solidität der Partner weitgehend durchsetzen.’ Anonymous, ‘Europas Regierungschefs...’, *Die Zeit* (1991).

<sup>66</sup> ‘Doch selbst in der längerfristigen Perspektive scheinen schon heute einige Länder chancenlos. Die auffallendsten Beispiele: Belgien, Griechenland und Italien haben so hohe Schulden, daß sie diese unmöglich in ein paar Jahren auf das geforderte Niveau abbauen können. [...] Nur geringe Chancen haben Portugal, Irland und Spanien. Sie müßten mächtig sparen und damit gewaltige soziale Spannungen riskieren’. Anonymous, ‘Die Richter in den roten Roben haben gesprochen...’, *Die Zeit* (1993). See also: Anonymous, ‘Die Turbulenzen des neuen Jahres...’, *Die Zeit* (1995).

<sup>67</sup> This is the title of: Anonymous, ‘Schreckgespenst Währungsunion’, *Die Zeit* (1995). For an overview of this pessimistic reporting, see the article by Ulrich Cartellieri, who himself comes to a more positive conclusion: Cartellieri, U., ‘Trotz Verfassungsgerichtsurteil, trotz politischer Willensbekundungen...’, *Die Zeit* (1993).

Before we turn to the temporal evolution of these discourses, it should be briefly noted that three of the topics listed in the table above feature word lists that are rather unclear and unprecise, mixing different themes and functions. While one of these topics could be loosely associated with *Rhetorical Terms* (topic 2), mostly related to political communication ('erklärung,' 'regierungen'), the other two topics were so diverse that they were simply labelled *Diverse Terms I* and *II* (topics 9 and 15). This mixing of diverse terms is not a failure of the model, or evidence of a general lack of identifiable topics, but is a rather common phenomenon in topic modelling analysis. For instance, topic models described by other researchers have repeatedly grouped mere linguistic patterns without useful meaning, clusters of methodological words, or even days of weeks into specific topics.<sup>68</sup> In his topic model analysis of the *Journal of Economic History*, Wehrheim finds topics that simply encompass technical expressions typical for quantitative methods.<sup>69</sup> Similarly, topic 5 (*Diverse Terms II*) features word such as 'bruttosozialprodukt' that necessarily appear throughout the decades in most newspapers in order to provide general descriptions of the state of the economy, but that on their own are not representative of a specific discourse. Instead of 'forcing' these lists of words in some vaguely labelled buckets, researchers usually ignore these unclear topics and focus on the ones that can be coherently labelled,<sup>70</sup> which is the approach that we have also followed in this paper.

The 15 topics estimated by the STM and described in detail above have different weights within the corpus, implying that the number of topics associated with the cultural, political and economic discourses does not necessarily reflect the respective discourse's importance. This issue can be addressed by visualising the expected proportion of the corpus that belongs to each topic, as plotted in the figure below. Each topic is identifiable via its label, with the length of the bar signalling its respective representation in the *Zeit* and thus its media significance. We see, for example, that *Football* (topic 4) is actually a relatively minor proportion of the newspaper discourse. The most common topics relate to *Tourism* (topic 12) and *Economic Competitiveness* (topic 6), reinforcing the notion that Germans' view of Europe's South was distinctively shaped first by experiences of travelling abroad and later by the

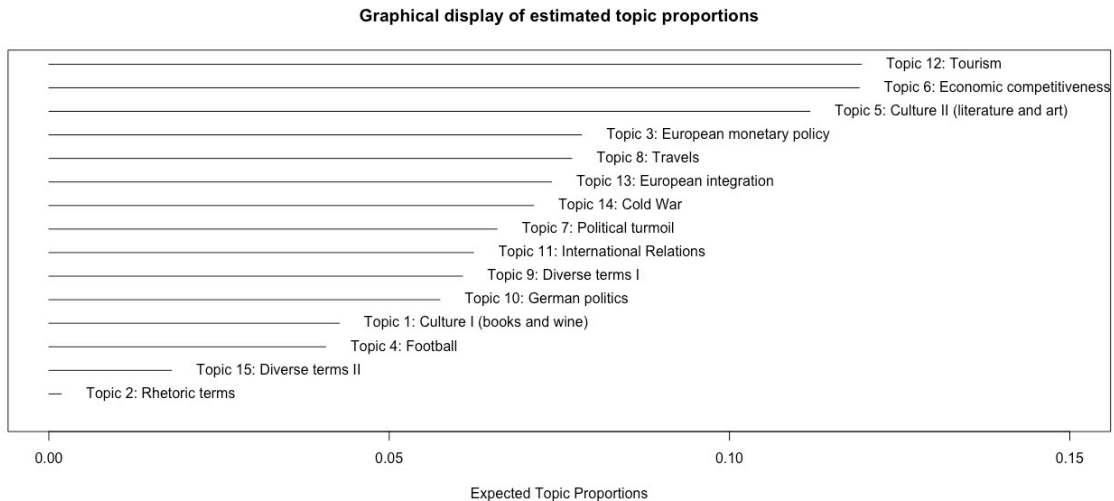
---

<sup>68</sup> Boyd-Graber, J., Mimno, D. and Newman, D.J., 'Care and feeding of topic model', in D.M. Blei, E.A. Erosheva, S.E. Fienberg and E.M. Airoldi (eds.), *Handbook of mixed membership models and their applications* (Boca Raton, 2015), p. 22.

<sup>69</sup> Wehrheim, L., 'Economic history goes digital: topic modeling the Journal of Economic History', *Cliometrica* 13 (2019), pp. 106f.

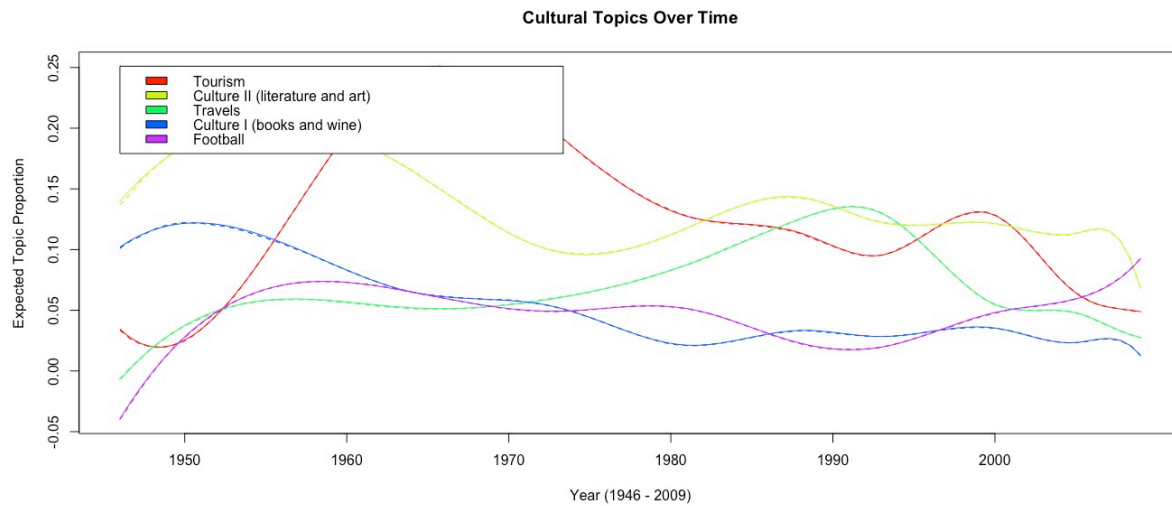
<sup>70</sup> See also the discussion in: Küsters, A., Volkind, L. and Wagner, A., 'Digital Humanities and the State of Legal History. A Text Mining Perspective', *Rechtsgeschichte – Legal History*, 27 (2019), forthcoming.

economic weaknesses of Southern European countries discussed during the process of European integration.



## RESULTS II: DEVELOPMENTS

After having gained a sufficient understanding of the topics' content, we can turn the relationship between these topics and the metadata coded during the construction process of the corpus (i.e. the variables *country* and *year*). These relationships can also play a key role in validating the STM's usefulness. After specifying the variable that one wishes to use for calculating an effect (all other variables are held at their sample median), one can plot this variable's effect on topics. We start by looking at the effects of the year variable on topical prevalence. The figure below plots the relationship between the respective year and the set of cultural topics, that is topics 12, 5, 8, 1, and 4, thereby visualising these topics' temporal dimension. Significantly, the *Tourism* topic (red line) peaks during the 1960s, that is, during the Golden Age of Germany's post-war reconstruction ('Wirtschaftswunder'), which allowed many Germans for the first time to travel to Southern Europe (via car) for an annual summer holiday. Still, the tourism topic remains relevant within the *Zeit* corpus and even increases its share of articles again around the year 2000.



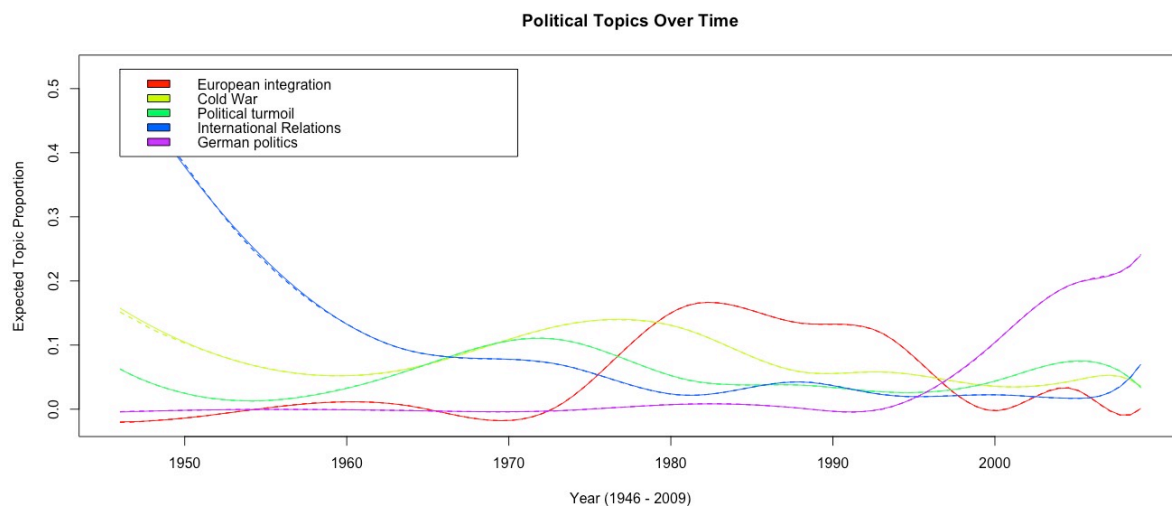
The STM therefore identifies *Tourism* as central cultural force in the media coverage of Spain. Such subcultural encounters and the resulting mutual relationships are by no means trivial. Tourism and identity are closely related: after exploring other countries the knowledge of the differences makes more present the characteristics of one's own country. This has repercussions on one's sense of national identity and how this identity might differ from others.<sup>71</sup> Especially in the process of constructing a post-war Europe, history and shared cultural heritage has played an important role in reinforcing new national identities and in national sentiment revival.<sup>72</sup> It is interesting to note that the change of the PIGS countries' image towards a more negative economic dimension was not caused by a decline in German tourism: with 13.5% of all holidays, Spain currently ranks first among the most popular foreign destinations for Germans, way ahead of Italy (7.7%) and Turkey (6.6%). By contrast, at 6%, the share of trips outside Europe is still quite low.<sup>73</sup>

<sup>71</sup> Lanfant, M.-F., Allcock, J.B. and Bruner, E.M., *International Tourism: Identity and Change* (London, 1995).

<sup>72</sup> About the importance of historical heritage as a political resource, see: Tunbridge, J.E. and Ashworth, G.J., *Dissonant Heritage: The Management of the Past as a Resource in Conflict* (Chichester, 1996). About post-communist tourism, see: Hall, D.R., 'Tourism change in Central and Eastern Europe', in A. Montanari and A.M. Williams (eds.), *European Tourism: Regions, Spaces and Restructuring* (London, 1995), pp. 221-44.

<sup>73</sup> Figures taken from: Mau, S. and Mewes, J., 'Transnationale soziale Beziehungen. Eine Kartographie der deutschen Bevölkerung', *Soziale Welt*, 58 (2007), p. 209.

Next, we turn towards the set of political topics, that is topics 13, 14, 7, 11, and 10, and estimate their development over time. The resulting figure, shown below, accurately resembles the shift in priorities from Cold War politics to European integration and finally to a media dominance of national political issues. The latter finding, i.e. the increased mixing of articles about Spain with discussions about German politics in the last years of our corpus, might indicate the increased relevance of discussions about PIGS countries in the German domestic political discourse during the Eurozone crisis, particularly during election times. The accompanying media discourse portrayed the crisis as one of ‘lazy’ Southern Europeans now punished for their ‘profligate’ lives, which justified austerity as the German government’s proposed crisis solution.<sup>74</sup> Commentators have particularly stressed the media pressure experienced by Merkel’s government due to important state elections that recurrently came up during the Eurozone crisis.<sup>75</sup>

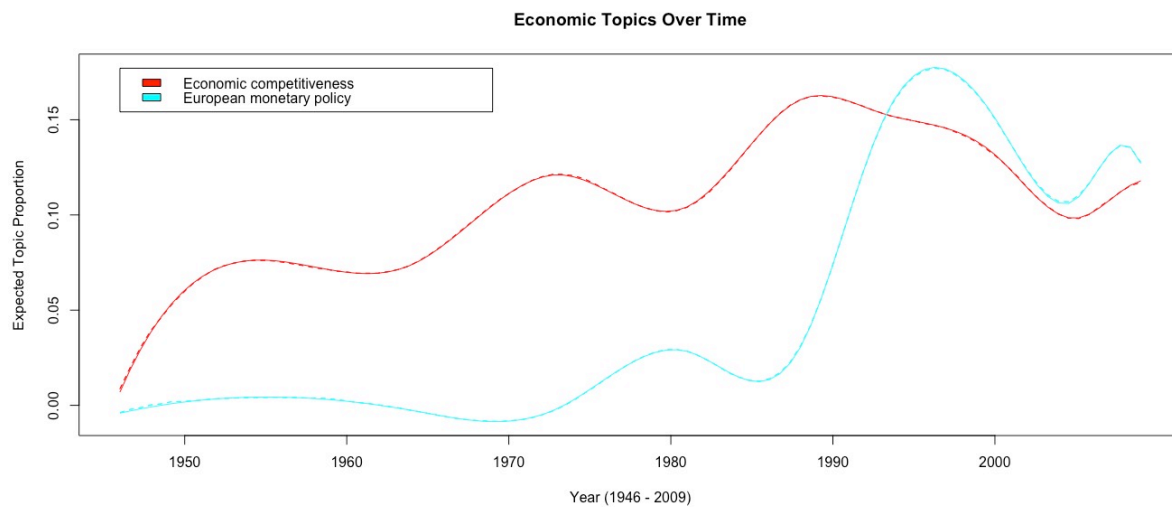


Finally, we turn to the economic discourse, which can be approximated with the estimated developments of topic 6 (*Economic competitiveness*) and topic 3 (*European monetary policy*). The resulting visualisation shows clearly that the initial focus on cultural and political topics in the *Zeit*'s reporting on Spain was only

<sup>74</sup> Petry, J., 'From PIIGS and the drive towards austerity: The discursive construction of the Eurozone crisis & its impact on European welfare states', *Joint Conference 'Drei-Länder-Tagung: Politik der Vielfalt'* (University of Innsbruck, 2013), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2329304](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2329304) (accessed: 06.09.2019).

<sup>75</sup> Mahnkopf, B., 'The euro crisis: German politics of blame and austerity – a neoliberal nightmare', *International Critical Thought*, 2/4 (2012), p. 480; Young, B. and Semmler, W., 'The European sovereign debt crisis: Is Germany to blame?', *German Politics and Society*, 29/1 (2011), p. 8.

recently replaced by economic issues. This trend can be explained with the deepening of European integration that started in the 1990s, culminating in the creation of the common currency area, and the heightened concerns over the price stability and economic competitiveness of the Southern European economies during this period. Thus, the figure indicates the point in time when the earlier PIGS heuristic, shaped by socio-cultural attributes as experienced for instance by German tourists, became supplemented with economic attributions in Germany's media discourse. This shift took roughly place between 1990 and 2000. The economic attributions inherent in this adapted country group heuristic in turn could be later moralised in the German media discourse accompanying the Eurozone crisis.<sup>76</sup>

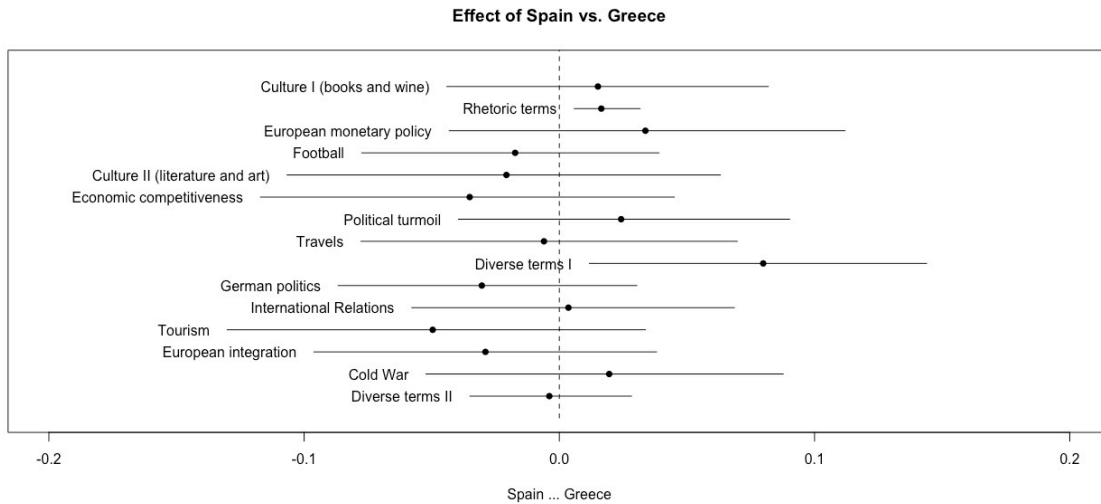


Besides analysing the topics' temporal dimensions via the *year* variable's effect on topic prevalence, we can also plot the influence of a topical content covariate. As mentioned, a topical content variable allows for the vocabulary used to write about a particular topic to vary. In our case, the STM has been fit with the *country* variable specified in the content option, meaning that we can calculate which words within a topic are more associated with a certain PIGS country versus another. In the figure below, we plot the change in topic proportions shifting from the default option, that is *Zeit* articles that focus exclusively on Spain, to articles that mention Spain *and* Greece. We see, for instance, that introducing Greece into the newspaper narrative implies that the journalist is, on average, less inclined to write about football, but more likely to write about political turmoil. Similar analyses can be done for the other PIGS countries,

<sup>76</sup> Fourcade, M., Steiner, P., Streeck, W. and Woll, C., 'Moral Categories in the Financial Crisis', *Socio-Economic Review*, 11/3 (2013), pp. 601-27.



thereby providing a convenient way for deepening our understanding of these countries' roles *within* the overall PIGS group heuristic – a topic that needs further investigation.



To sum up, estimating a topic model such as the STM presented in this paper has three crucial advantages, especially when compared to classic close reading of texts. First, it is an objective method which does not require the imposition of pre-defined categories. This is especially helpful when dealing with a contested, emotional issue such as the PIGS country group heuristic: by avoiding the ex-ante imposition of a certain picture or definition of the PIGS countries, the researcher can leave it to the sources, that is the corpus, to indicate which words have been associated empirically with this country. Secondly, the estimated topics are explicit, so other researchers can reproduce the analysis. Thirdly, the computational power allows to understand and structure large corpuses of texts, thereby facilitating the complementing qualitative analysis via traditional close reading. In this case, STM allowed us to survey 2,443 articles published between 1946 and 2009 in the German newspaper *Die Zeit*. As has become clear, these quantitative techniques still depend on the researcher's judgment and thus should be seen as a helpful complement, and not substitute, in historic research.

## CONCLUSION

The onset of the Eurozone crisis popularised the PIGS acronym, conveying allegedly 'lax' Southern European attitudes towards inflation and state deficit and contrasting Portugal, Italy, Greece, and Spain

unfavourably with their Northern European neighbours. In order to investigate the historic formation of this country group heuristic, we apply a Structural Topic model (STM) to all 2,443 articles published between 1946 and 2009 in the German newspaper Die Zeit that mention Spain. The estimated topics and their evolution over time confirm a strong association of Spain with other Southern European countries as an ongoing characteristic of the German media discourse for more than sixty years. Moreover, the STM allows us to distinguish and characterise a cultural, a political and an economic dimension of the PIGS countries' picture in the German media. We find that the initial PIGS heuristic shaped by socio-cultural attributes, that mainly reflected the experiences of German tourists, was later supplemented with economic ascriptions, which in turn provided the basis for the moralising media discourse accompanying the Eurozone crisis. It is up to future research to complement this picture by including other newspapers or even other types of media in the analysis.

## REFERENCES

- Akaliyski, P., 'United in diversity? The convergence of cultural values among EU member states and candidates', *European Journal of Political Research*, 58/2 (2018), pp. 388-411.
- Anonymous, 'Hail, the Swabian housewife: The German mentality', *The Economist*, 410/8872 (2014), pp. 45-46.
- Arora, S., Ge, R., Halpern, Y., Mimno, D., Moitra, A., Sontag, D., Wu, Y. and Zhu, M., 'A Practical Algorithm for Topic modeling with Provable Guarantees', in S. Dasgupta and D. McAllester (eds.), *Proceedings of the 30th International Conference on Machine Learning*, 28 (2013), pp. 280-288.
- Blei, D.M., 'Topic modeling and Digital Humanities', *Journal of Digital Humanities*, 2/1 (2012), <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/> (accessed: 06.09.2019).
- Blei, D.M., Ng, A., Jordan, M., 'Latent Dirichlet Allocation', *Journal of Machine Learning Research*, 3 (2003), pp. 993-1022. These models were subsequently extended, see e.g. the Correlated Topic model: Blei, D.M. and Lafferty, J.D., 'A correlated topic model of Science', *The Annals of Applied Statistics*, 1/1 (2007), pp. 17-35.

Blyth, M., 'Structures Do Not Come with an Instruction Sheet: Interests, Ideas, and Progress in Political Science', *Perspectives on Politics*, 1/4 (2003), pp. 695-706.

Bonilla, T. and Grimmer, J., 'Elevated threat levels and decreased expectations: how democracy handles terrorist threats', *Poetics*, 41/6 (2013), pp. 650-669.

Boyd-Graber, J., Mimno, D. and Newman, D.J., 'Care and feeding of topic model', in D.M. Blei, E.A. Erosheva, S.E. Fienberg and E.M. Airoldi (eds.), *Handbook of mixed membership models and their applications* (Boca Raton, 2015), pp. 3-41.

Brazys, S. and Hardiman, N., 'The 'PIIGS' acronym had a clear negative impact on the response of financial markets to the "PIIGS countries" during the crisis', *LSE Blog Post* (12.12.2014), <https://blogs.lse.ac.uk/euoppblog/2014/12/12/the-piigs-acronym-had-a-clear-negative-impact-on-the-market-treatment-of-the-piigs-countries-during-the-crisis/> (accessed: 29.08.2019).

Breuer, T, 'Ein Dauerplatz an der Sonne', *Praxis Geographie*, 3 (2002), pp. 21-27.

Brunnermeier, M.K., James, H. and Landau, J.-P., *The Euro and the Battle of Ideas* (Princeton, 2016).

Dainotto, R., *Europe. In Theory* (Durham, 2007).

DiMaggio, P., Nag, M. and Blei, D., 'Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. Government arts funding', *Poetics*, 41/6 (2013), pp. 570-606.

Elliott, J., *Europe Divided* (London, 2000).

European Council and Committee for the Study of Economic and Monetary Union, *Report on economic and monetary union in the European Community* (Unipub, 1989).

Ferri, P., Lusiani, M. and Pareschi, L., 'Accounting for Accounting History: A Topic modeling Approach (1996–2015)', *Accounting History*, 23/1-2 (2018), pp. 173-205.

Fourcade, M., Steiner, P., Streeck, W. and Woll, C., 'Moral Categories in the Financial Crisis', *Socio-Economic*

Franks, J., Barkbu, B., Blavy, R., Oman, W. and Schoelermann, H., 'Economic Convergence in the Euro Area: Coming Together or Drifting Apart?', *IMF Working Paper* (January 2018), <https://www.imf.org/en/Publications/WP/Issues/2018/01/23/Economic-Convergence-in-the-Euro-Area-Coming-Together-or-Drifting-Apart-45575> (accessed: 06.09.2019).

Garrido, E., 'Exploring North-South Identities Using NLP: The Image of Spain in the German Weekly Die Zeit', *Bulletin for Spanish and Portuguese Historical Studies*, 43/1 (2018), pp. 53-69.

Glorius, B. and Domínguez-Mujica, J., *European Mobility in Times of Crisis: The New Context of European South-North Migration* (Bielefeld, 2017).

Gros, D., 'How Fit are the Candidates for EMU?', *The World Economy*, 23/10 (2000), pp. 1367-1377.

Hall, D.R., 'Tourism change in Central and Eastern Europe', in A. Montanari and A.M. Williams (eds.), *European Tourism: Regions, Spaces and Restructuring* (London, 1995), pp. 221-44.

Hayo, B. and Neumeier, F., 'The social context for German economists: public attitudes towards macroeconomic policy in Germany', in G. Bratsiotis and D. Cobham (eds), *German Macro: How It's Different and Why That Matters* (Policy File, 2016), pp. 64-72.

Hien, J., 'The Religious Foundations of the European Crisis', *JCMS: Journal of Common Market Studies*, 57/2 (2019), pp. 185-204.

Jacobi, C., van Atteveldt, W. and Welbers, K., 'Quantitative analysis of large amounts of journalistic texts using topic modeling', *Digit J*, 4/1 (2015), pp. 89-106.

James, H., 'Rule Germania', in T. Beck and H.-H. Kotz (eds), *Ordoliberalism: A German oddity?* (London, 2017), pp. 25-30.

Kordel, S., 'Lifestyle-Mobilitäten deutscher Senioren in Spanien. Das Beispiel der Gemeinde Torrox an der Costa del Sol', *Mitteilungen der Fränkischen Geographischen Gesellschaft*, 58 (2011), pp. 53-66.

Küsters, A., Volkind, L. and Wagner, A., 'Digital Humanities and the State of Legal History. A Text Mining Perspective', *Rechtsgeschichte – Legal History*, 27 (2019), forthcoming.

Lafleur, J., and Stanek, M. (eds.), *South-North Migration of EU Citizens in Times of Crisis* (Dordrecht, 2017).

Lanfant, M.-F., Allcock, J.B. and Bruner, E.M., *International Tourism: Identity and Change* (London, 1995).

Lucas, C., Nielsen, R., Roberts, M., Stewart, B., Storer, A., Tingley, D., 'Computer Assisted Text Analysis for Comparative Politics', *Political Analysis*, 23/2 (2015), pp. 254-277.

Mahnkopf, B., 'The euro crisis: German politics of blame and austerity – a neoliberal nightmare', *International Critical Thought*, 2/4 (2012), p. 472-485.

Mau, S. and Mewes, J., 'Transnationale soziale Beziehungen. Eine Kartographie der deutschen Bevölkerung', *Soziale Welt*, 58 (2007), pp. 203-222.

Mee, S., 'Monetary mythology: The West German central bank and historical narratives, 1948-78' (unpublished D.Phil. thesis, University of Oxford, 2016).

Mundell, R., 'Currency Areas, Common Currencies and EMU', *The American Economic Review*, 87/2 (1997), pp. 214-216.

Nelson, R.K., 'Mining the Dispatch', *Blog*: <http://dsl.richmond.edu/dispatch/Topics/view/28> (accessed: 16.11.2017).

Newman D.J. and Block, S., 'Probabilistic topic decomposition of an eighteen-century American newspaper', *J Am Soc Inform Sci Technol*, 57/6 (2006), pp. 753-767.

Petry, J., 'From PIIGS and the drive towards austerity: The discursive construction of the Eurozone crisis & its impact on European welfare states', *Joint Conference 'Drei-Länder-Tagung: Politik der Vielfalt'* (University of Innsbruck, 2013), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2329304](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2329304) (accessed: 06.09.2019).

*Review*, 11/3 (2013), pp. 601-27.

Roberts, M.E., Stewart, B.M. and Airoldi, E., 'A model of text for experimentation in the social sciences', *Journal of the American Statistical Association*, 111/515 (2016), pp. 988-1003.

Roberts, M.E., Stewart, B.M. and Tingley, D., 'stm: R Package for Structural Topic models', *Journal of Statistical Software*, available online: <https://cran.r-project.org/web/packages/stm/vignettes/stmVignette.pdf> (accessed: 06.09.2019).

Shambaugh, J., 'The Euro's Three Crises', *Brookings Papers on Economic Activity*, 43/1 (2012), pp. 157-231.

Silge, J. and Robinson, D., *Text Mining with R: A Tidy Approach* (eBook, 2017), <https://www.tidytextmining.com> (accessed: 06.09.2019).

Tokarski, P., 'Divergence and Diversity in the Euro Area. The Case of Germany, France and Italy', *SWP Research Paper*, 06 (2019), <https://www.swp-berlin.org/10.18449/2019RP06/> (accessed: 06.09.2019).

Tunbridge, J.E. and Ashworth, G.J., *Dissonant Heritage: The Management of the Past as a Resource in Conflict* (Chichester, 1996).

Wehrheim, L., 'Economic history goes digital: topic modeling the Journal of Economic History', *Cliometrica* 13 (2019), pp. 83-125.

Young, B. and Semmler, W., 'The European sovereign debt crisis: Is Germany to blame?', *German Politics and Society*, 29/1 (2011), pp. 1-24.

Zamora-Kapoor, A., and Coller, X., 'The Effects of the Crisis: Why Southern Europe?', *American Behavioral Scientist*, 58/12 (2014), pp. 1511–1516.

Zimmerman, K.F., 'European Migration: Push and Pull', *International Regional Science Review*, 19/1–2 (1996), pp. 95-128.

## Acknowledgements

This project has received funding from the H2020-EU.3.6 – SOCIETAL CHALLENGES – Europe in a Changing World – Inclusive, Innovative and Reflective Societies under grant agreement no. 649307. The project UPIER is financially supported by the HERA Joint Research Programme ([www.heranet.info](http://www.heranet.info)) which is co-funded by AHRC, AKA, PT-DLR, CAS, CNR, DASTI, ETAG, FCT, FNR, F.R.S.-FNRS, FWF, FWO, HAZU, IRC, LMT, MIZS, MINECO, NCN, NOW, RANNÍS, RCN, SNF, VIAA, VR and The European Community, SOCIETAL CHALLENGES – Europe in a Changing World – Inclusive, Innovative and Reflective Societies under grant agreement no. 649307.



Published by the UPIER project

**Address** University of Oxford, Faculty of History, George Street, Oxford, OX1 2RL, UK

**Contact** [UPIER@history.ox.ac.uk](mailto:UPIER@history.ox.ac.uk)

**Website** [www.upier.web.ox.ac.uk](http://www.upier.web.ox.ac.uk)

Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the UPIER project or the authors.