The association between press coverage of the economy and fertility in the United Kingdom

Paolomaria Fabrizio[a,c], Ann Berrington[b], Raffaele Guetto[a], Jason Hilton[b], Daniele Vignoli[a]

[a] *University of Florence, Florence, Italy*
[b] *University of Southampton, Southampton, UK*
[c] *Corresponding author: paolomaria.fabrizio@unifi.it*

Abstract

In today's globalized world, media plays a crucial role in shaping perceptions and expectations about states of the economy, therefore potentially affecting fertility behaviors. This paper aims to analyse the newspaper coverage of economic narratives in the UK over the past 20 years and assess their impact on general fertility trends. We use data from the LexisNexis database to access the UK press and classify economic news articles with a sentiment through machine learning techniques, ultimately developing press-based indices. We employ an ARDL model to test the effects of our press-based indices on national fertility trends.

Keywords: Media narratives of the economy, Fertility, UK

1. Introduction and Aim

In the wake of the Great Recession, the vast majority of European countries experienced a significant decrease in fertility rates, often irrespective of their pre-existing fertility levels. This declining pattern presents significant challenges to the current demographic understanding. Studies suggest that the decline in contemporary European fertility rates cannot be entirely explained through traditional economic indicators. Because fertility choices involve long-term considerations and it is increasingly difficult to predict their consequences on parents' life trajectory, uncertainty plays a significant role (Vignoli et al., 2020).

The growing economic uncertainty in contemporary societies, particularly in the Western world, emphasizes the role of media narratives about the economy in shaping people's fertility decisions. This influence may happen due to the media's selection of particular topics and their presentation of information from specific perspectives and in certain tones. By examining the association between how the economy is portrayed in news media and people's choices regarding their reproductive behaviour, it is possible to explore the connection between media narratives and fertility (Guetto et al., 2023).

1.1. Objective of research and research questions

This paper aims to study the association between an aggregate measure of fertility and the media coverage of the economy in the UK.

Building upon this premise, the present work aims to address the following research questions:

- To what extent does the volume of the economic media coverage influence fertility trends in the UK?
- Does this association change based on the sentiment of the economic news?

To answer these research questions, we will utilize a newspaper dataset, LexisNexis, to construct two press-based indices on economic news, exploiting a machine learning approach: a coverage-based index and a sentiment-based index.

Furthermore, to examine the association between the press coverage and fertility, an ARDL model, will be used. In addition, several related variables that could potentially confound the observed associations will be considered. These variables include socioeconomic indicators and relevant events or policy changes that could have impacted both press coverage and fertility trends.

2. Literature review

Studies conducted in Italy (Guetto et al.2021, 2022) have recently delved into the effects of COVID-19-induced uncertainty. By exposing participants to simulated news bulletins

regarding the expected duration of the pandemic, these investigations revealed a cause-and-effect relationship between narratives concerning the pandemic and individuals' union and fertility intentions.

Vignoli et al., 2022 performed a controlled laboratory experiment in Florence and Oslo, where they simulated news bulletins regarding the economic outlook of the country. The findings revealed a causal impact of economic narratives on individuals' intentions regarding fertility. In a separate study, Comolli and Vignoli, 2021 estimated a regression discontinuity design during Italy's debt crisis in November 2012. Their estimates indicated a decrease in birth rates due to an increase in perceived economic uncertainty. Although these studies indicate the potential impact of media on fertility behaviours, they have not specifically examined the direct effects of the phenomenon. To the best of our knowledge, only Schneider, 2015 and Guetto et al., 2023 have examined the association between economic news coverage and fertility. The former paper found evidence that press coverage of the economic recession contributed to the decline in state-level fertility rates in the United States, independent of traditional economic measures while in the latter paper, the authors merged individual-level data from the data on the sentiment of economic news reported by the evening newscast of the Italian TV channel one. Their analysis shows that fertility behaviour, net of traditional economic indicators, is associated with both the incidence and tone of news coverage on the state of the economy, particularly a deterioration in the tone of economic news discourages fertility among young adults without children, while an improvement in the tone of media news fosters individual fertility.

3. Data and methods

3.1. LexisNexis

LexisNexis is a unique database containing a comprehensive collection of news articles from various newspapers from different European countries. The collection related to the United Kingdom presents over 70 million articles from 1970 to 2020. The data provided by LexisNexis offers a diverse array of news originating from multiple sources and media types, including national newspapers, local newspapers, online newspapers, and news agencies. To construct our indices we did not consider articles of little interest to non-specialists (e.g. market or regulatory information). Such articles do not have a wide readership and are unlikely to affect fertility making decisions. Furthermore, we did not consider local newspapers since the coverage is not uniformly distributed across the entire territory of the United Kingdom. For these reasons, we removed news agencies and local newspapers from the analysis.

3.2. Indices based on newspaper media

The newspapers' articles have been used to create an index that measures the volume of the monthly economic news coverage and another index that assesses the sentiment of this news. To construct the aforementioned indices, we first developed guidelines for manually annotating the data. Next, we extracted a sample of articles from the Lexis Nexis database,

which were then annotated according to these guidelines. Finally, the annotated sample was used to train a machine learning model, the Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2019), to predict annotations for the full dataset.

These guidelines included rules for distinguishing between economic and non-economic articles, and for classifying economic articles as negative, positive, or neutral. The classification tasks involved principally two steps: the first consisted of the pre-training of a BERT model on English text corpora and the second included the fine tuning phase. The pre-training model used is *bert-base-cased*, which has been trained on BookCorpus, a dataset consisting of 11,038 unpublished books and English Wikipedia (excluding lists, tables and headers). It has been made it available on the Hugging Face platform (https://huggingface.co/bert-base-cased), a resource for sharing machine learning models. The first index we will construct, the **Economic Coverage Index** (ECI), follows the Economic Policy Uncertainty index construction of Baker et al., 2016: first we count, for each month of analysis, the proportion of articles in a given month that the BERT model classifies as economic. These raw counts are not a precise measure and should be contextualized with respect to the newspaper, time period, and other news written by the same newspaper during that time. For this purpose we divided by the total number of articles in the same newspaper and month: $(\frac{\#\text{articles talking about economy}}{\#\text{total articles}})$; then we standardized each newspaper time series to unit standard deviation and averaged. The second index, the **Economic Sentiment Index** (ESI), is built according to the same approach but, in this case, we count, for each month of analysis, the articles in a given month that the BERT model classifies as positive minus the ones classified as negative scaled by the sum of the positive plus the negative ones $\frac{(\#\text{pos}-\#\text{neg})}{(\#\text{pos}+\#\text{neg})}$, excluding the neutral ones.

Negative and positive sentiment are both considered in a unique index but neutral news is excluded to simplify the index calculation and more effectively highlight variation between positive and negative sentiments (Guetto et al., 2023; Soo, 2018).

Furthermore, we weight these indices using circulation figures data from the ABC website https://www.abc.org.uk/ to take into account the size of the audience that is potentially exposed to certain news. Data for printed newspapers is available for the entire period under consideration, while for online newspapers, data is only available until 2017. It is assumed that for the remaining period, the values for online newspapers remain constant, assuming the same value as the last available months.

3.3. Statistical analysis

The ARDL model [1] is well-suited for examining data collected over time, allowing the exploration of potential correlations and quantifying the magnitude of the association between time series variables. This model captures the lagged effects of the dependent variable and the lagged effects of each variable on the response. It allows for understanding how changes

[1] $y_t = c_0 + c_1 t + \sum_{i=1}^{p} b_{y,i} y_{t-i} + \sum_{j=1}^{k} \sum_{l=0}^{q_j} b_{j,l} x_{j,t-l} + \epsilon_t$

in one variable impact the dependent variable over time. Therefore, we will incorporate several relevant variables, such as socioeconomic indicators and significant events, to control for potential confounding factors.

The **dependent variable** of the model is the monthly general fertility. Monthly births are taken from the different national statistical agencies: ONS, (England and Wales), NRS (Scotland) and NISRA (Northern Ireland). The denominator was retrieved from the ONS site in the dataset "Mid-2001 to mid-2020 detailed time series edition of this dataset".

The covariates to be considered in the model are those related to macroeconomic indicators and the constructed media indicators from 2000 to 2020:

- the chained volume of gross value added estimate of monthlyGDP (CVGVA GDP);
- the seasonally adjusted unemployment rate (aged 16 and over) (UR);
- the measure of the monthly consumer price inflation(CPI);
- the average weekly earnings of the Great Britain (EARN);
- the monthly series of the press economic coverage (ECI)
- the monthly series of the sentiment related to economic news (ESI)

4. Preliminary results and expected findings

The BERT model for the ECI was trained with 2000 newspapers' abstracts, of which 1385 did not refer to the economy and 615 were related to economic themes.

The training set was split into a 75/25 ratio for training and testing purposes. Afterwards, the test set was further divided into a 50/50 ratio to create a validation set and a final test set. Finally, the training process returned an f1-score of 84%.

The table below (Table 1) gives the precision, and f1 scores of the model for the final, indicating the model performs reasonably well in differentiating economic from non-economic news.

	precision	**recall**	**f1-score**
OffTopic	0.84	0.94	0.89
InTopic	0.84	0.65	0.73
Total			0.84

Table 1: Model In-OffTopic performances

The sentiment model was trained with a larger training sample since there are three classification labels: positive, neutral and positive.

It was made of 5476 newspapers' economic abstracts, of which 1938 had a negative sentiment, 1390 were neutral and 2148 had a positive sentiment.

The table below (Table 2) gives the precision, and f1 scores of the sentiment model.

	precision	**recall**	**f1-score**
Negative	0.78	0.79	0.79
Neutral	0.35	0.38	0.36
Positive	0.72	0.68	0.70
Total			0.66

Table 2: Sentiment Model performances

To build the indices we adopt a selection strategy in order to include the most widespread printed and online newspapers. The newspapers featuring in the list of those with the largest circulation during either 2000 or 2020. The printed newspapers selected are the "Daily Express", the "Daily Mail", the "Daily Mirror", the "Financial Times", the "Sunday Express", the "Sunday Mail", the "Sunday Mirror", the "Guardian" and the "Observer". The online newspapers are the "Daily Star Online", "Express Online", "The Guardian Online", "The Mail Online", "The Mirror Online", "The Post Online" and "The Telegraph Online".

Figure 1 shows the ECI and the ESI lagged by 12 months and the dependent variable.

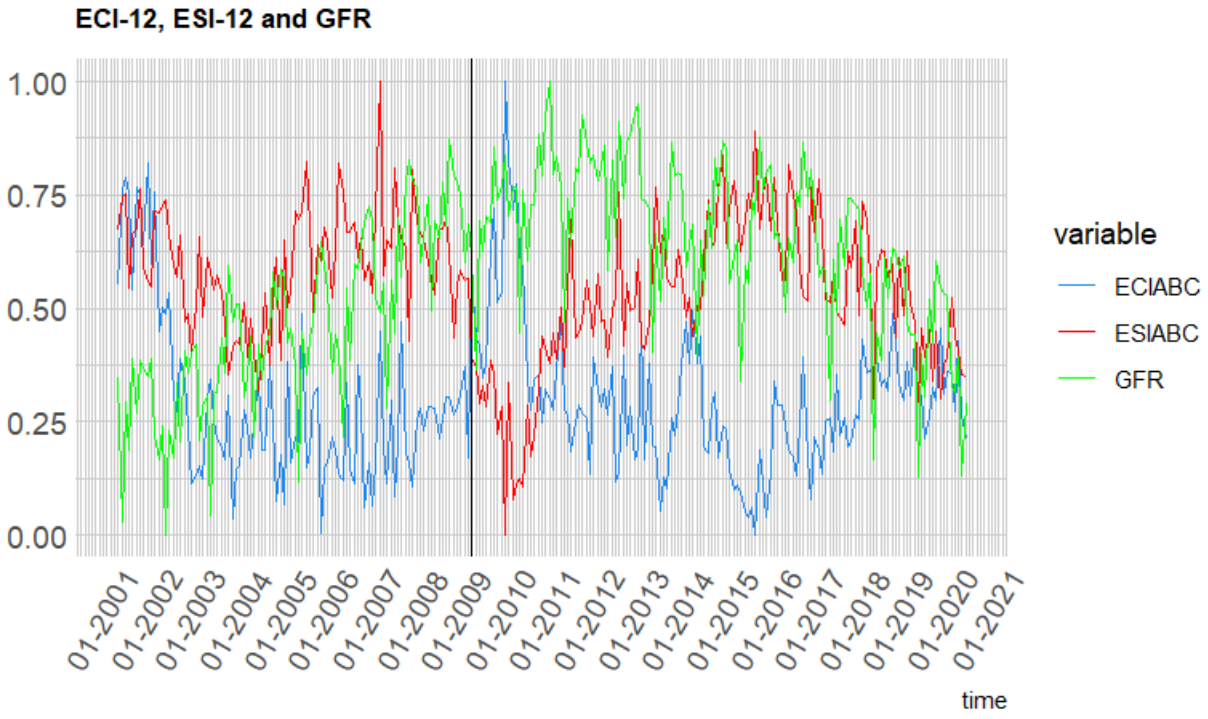

Figure 1: Lagged ECI, ESI and GFR

The ECI reveals that the newspapers appeared to allocate more space to economics during and after the crisis period while the ESI has had a strong decline during the crisis followed by a rise in the post-crisis period and then fell again with the arrival of covid. Moreover, after 2012 the GFR and the ESI tend to move in the same direction. There appears to be an inverse relationship between the ECI and the ESI, where a high ECI corresponds with a low ESI, and vice versa.

Since we are attempting to explain the overall fertility patterns, it is necessary to lag the covariates, as individuals require time to comprehend the underlying phenomena and make decisions. Additionally, the consideration of the nine-month gestation period is crucial. For this reason, we lagged covariates by 12 months [2].

We estimated an ARDL model, choosing the lag orders of the covariates through the BIC selection criterion. To account for the strong seasonality of the GFR, we included

[2]The lag period was chosen in reference to Schneider, 2015, who included a one-year lagged measure of state economic conditions in the model and also because after trying various lags, the Bayesian Information Criterion (BIC) consistently preferred the 12-month lag.

period dummy variables, along with a crisis dummy variable and its interaction with the journalistic indices. In Table 3 we show the results of the model.

Table 3: ARDL results

	Dependent variable: GFR
L(GFR, 1)	0.535^{***}
	(0.064)
L(GFR, 2)	0.183^{***}
	(0.064)
CVGVA GDP	-0.004
	(0.012)
L(CVGVA GDP, 1)	-0.042^{***}
	(0.013)
L(CVGVA GDP, 2)	0.042^{***}
	(0.013)
UR	-0.039^{**}
	(0.016)
L(UR, 1)	0.049^{***}
	(0.016)
CPI	-0.008
	(0.007)
EARN	0.019^{***}
	(0.004)
ECI	0.020
	(0.022)
ESI	-0.057^{***}
	(0.022)
feb	-0.430^{***}
	(0.023)
mar	0.186^{***}
	(0.036)
apr	-0.085^{**}
	(0.033)
mag	0.177^{***}
	(0.025)
giu	-0.044^{*}
	(0.026)
lug	0.197^{***}

Continued on next page

	Dependent variable: GFR
	(0.024)
ago	0.025
	(0.027)
sep	0.042
	(0.027)
oct	0.045*
	(0.025)
nov	−0.250***
	(0.025)
dic	−0.052*
	(0.029)
varCris	−0.008
	(0.028)
ECI:varCris	0.034
	(0.035)
ESI:varCris	0.129***
	(0.032)
Constant	−0.249
	(0.179)
Observations	228
R^2	0.959
Adjusted R^2	0.954
Residual Std. Error	0.067 (df = 202)
F Statistic	189.016*** (df = 25; 202)
Note:	*p<0.1; **p<0.05; ***p<0.01

The results indicate several significant coefficients in our ARDL model. Notably, the lagged values of the GFR are highly significant, suggesting a strong autoregressive component in the fertility dynamics.

In the ARDL framework, short-term effects [3] capture immediate changes in the dependent variable in response to explanatory variables, while long-term effects [4] represent cumulative variations over time, reflecting the impact of explanatory variables on the dependent variable in the long run.

In the short-run period the UR and the ESI (without interaction) have a negative significant effect on GFR while the earnings and the ESI with the interaction, have the opposite effect.

In the long-run period the EARN, the ESI and the ESI with interactions have significant effects. In particular, the ESI (with no interaction) has a small negative effect on the GFR while the EARN and the ESI with interaction have a positive effect on the response variables.

To summarize, it seems that after the economic crisis in 2009, media play a significant role alongside economic indices in influencing individual choices at the macro level

To summarize, after the economic crisis, it seems that the media, against economic indices, have a significant influence on individual decisions. Specifically, at the macro level, the manner in which news is conveyed (as measured by the ESI) could substantially affect fertility.

Variables related to economic sentiment and news coverage, such as ESI (economic sentiment index) and ECI (economic coverage index), also display significant effects, indicating the influence of media perception on fertility behavior. Additionally, the inclusion of interaction terms between economic indicators and crisis dummies suggests that the impact of economic factors on fertility may vary during periods of crisis.

5. Conclusion

This study is an attempt to study the association between newspaper media narratives about the economy and fertility trends in the UK. Although this association remains largely unexplored within the context of the UK, existing literature suggests the existence of this relationship in other countries (Guetto et al., 2023, Schneider, 2015).

This work highlights a selection strategy which involves selecting and weighting newspapers based on their circulation and readership, resulting from the need to capture a representative and impactful coverage of the phenomenon for a more comprehensive analysis.

Our study advances the existing literature through our novel approach to classifying news articles and in the way we weight the indices by circulation figures. This could potentially pave the way for exploring the interplay between press and fertility trends at a macro level, serving as a model for future research in different countries.

---

[3] $\frac{\partial x_{j,t}}{\partial y_t} = b_{j,0}, \quad j \in \{1, \ldots, k\}$

[4] $\frac{\partial y_{t+\infty}}{\partial y_t} = \theta_j = \frac{\sum_{l=0}^{q_j} b_{j,l}}{(1-\sum_{i=1}^{p} b_{y,i})}, \quad j \in \{1, \ldots, k\}$

References

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty*. *The Quarterly Journal of Economics*, *131*(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

Comolli, C. L., & Vignoli, D. (2021). Spreading uncertainty, shrinking birth rates: A natural experiment for italy. *European Sociological Review*, *37*(4), 555–570.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding.

Guetto, R., Bazzani, G., Vignoli, D., et al. (2022). Narratives of the future and fertility decision-making in uncertain times. an application to the covid-19 pandemic. *Vienna Yearbook of Population Research*, 1–38.

Guetto, R., Morabito, M. F., Vollbracht, M., & Vignoli, D. (2023). Fertility and Media Narratives of the Economy: Evidence From Italian News Coverage. *Demography*, *60*(2), 607–630. https://doi.org/10.1215/00703370-10607928

Guetto, R., Vignoli, D., & Bazzani, G. (2021). Marriage and cohabitation under uncertainty: The role of narratives of the future during the covid-19 pandemic. *European Societies*, *23*(sup1), S674–S688.

Schneider, D. (2015). The great recession, fertility, and uncertainty: Evidence from the united states. *Journal of Marriage and Family*, *77*(5), 1144–1156.

Soo, C. K. (2018). Quantifying Sentiment with News Media across Local Housing Markets. *The Review of Financial Studies*, *31*(10), 3689–3719. https://doi.org/10.1093/rfs/hhy036

Vignoli, D., Bazzani, G., Guetto, R., Minello, A., & Pirani, E. (2020). Uncertainty and narratives of the future: A theoretical framework for contemporary fertility. In R. Schoen (Ed.), *Analyzing contemporary fertility* (pp. 25–47). Springer International Publishing. https://doi.org/10.1007/978-3-030-48519-1_3

Vignoli, D., Minello, A., Bazzani, G., Matera, C., & Rapallini, C. (2022). Narratives of the future affect fertility: Evidence from a laboratory experiment. *European Journal of Population*, *38*(1), 93–124.