PERCEIVE DEL. 3.4:

Horizon 2020
European Union Funding
for Research & Innovation

PERCEIVE

Perception and Evaluation of Regional and Cohesion Policies
by Europeans and Identification with the Values of Europe

Perception and Evaluation of Regional and Cohesion Policies by Europeans and Identification with the Values of Europe

PERCEIVE

GA nr. 693529

D 3.4 'Report on the statistical analysis of communication efforts' impact and effectiveness on citizens' awareness and appreciation of EU-financed projects'

AUTHORS OF THIS REPORT AND EDITORS OF THE DELIVERABLE:

- Vitaliano Barberio, Research Institute for Urban Management and Governance, Vienna University of Economics and Business, vitaliano.barberio@wu.ac.at
- Ines Kuric, Research Institute for Urban Management and Governance, Vienna University of Economics and Business, ines.kuric@wu.ac.at
- Edoardo Mollona, Università di Bologna, edoardo.mollona@unibo.it
- Luca Pareschi, Università di Bologna, luca.pareschi@unibo.it

LEADING PARTNER:
Wirtschaftsuniversität Wien [WU, Austria]

ESTIMATED PERSON/MONTHS PER PARTNER:
Wirtschaftsuniversität Wien [WU, Austria]: 6
Alma Mater Studiorum – Università di Bologna [UNIBO, Italy]: 5

SUBMISSION DATE: *3nd of July, 2018*

Table of Contents

0. Summary

0.1 Introduction

Why is a focus on policy communication relevant to the practice and study of European integration? Over the two last two decades, no clear pattern seems to have emerged linking the efforts of the European Union (EU) to foster local development and reduction of disparities with a generalized increase in the levels of European identity. This constitutes a relevant challenge for both policy makers and local implementers as well as for students of the European integration process. In this context, it is surprising how little attention has been paid to the potential role of cohesion policy – the EU's largest investment policy aimed at promoting regional development and reducing disparities.

Based on a recognition of previous research and direct experience acquired through interaction with practitioners of the cohesion policy field, in this report, we point out the importance of 'awareness' and 'appreciation' (of EU efforts) by EU citizens. This means that, albeit not a general rule, those who are aware of the policy often appreciate EU developmental policies and, when these two conditions are matched with each other, tend to fell more European. In managerial terms, this gap between institutions and citizens points to the centrality of policy communication. That is, we need to better understand the mechanisms and outcomes of communicating policy in todays' practice of policy implementation.

Accordingly, this report aims at increasing our understanding of a number of strategic aspects connected with the task of communicating EU regional cohesion policy at the local level. In more detail, two main aspects are dealt with: First, the extent to which local communication activities take part in a wider international discourse (i.e. through topics). Here, we expand our previous work (i.e. see PERCEIVE deliverables 5.3 and 3.3) on communicating EU cohesion policy through social media by analyzing contents generated by Local Managing Authorities' (hereafter referred to as LMAs) use of Facebook. And second, the impact and effectiveness of communication efforts (i.e. the effect of the amount of communication investments and the content of communication activities on citizens' awareness of EU policies and appreciation of EU actions). Inquiring this aspect entailed a statistical analysis of the association between communication efforts which LMAs disseminate the aims and accomplishments of regional cohesion policy on the one hand, and awareness as well as appreciation of the policy by EU citizens on the other hand. The outcomes of the social media analysis described as the first aspect above are also included as variables of the statistical analysis of the second focal aspect.

0.2 Main findings

The main findings of our analyses are:

The amount of **pro capita structural funding** significantly and positively correlates with **awareness** of the policy, **perceived personal benefit** and **political appreciation** of countries' membership in the EU. This result is stable across several model specifications including different combinations of variables.

The amount of **pro capita communication investment in the current period (2014-20)** related to regional operative plans significantly and positively correlates with the probability for citizens to perceive a **personal benefit** deriving from EU policy. This interrelation has an opposite effect for the case of political appreciation (i.e. membership in the EU). This finding suggests non-linearity in the effects of communication (i.e. perceived personal benefit does not automatically translate into higher level political appreciation), and calls for further investigation at the current state.

The **pro capita communication investment in the past period (2007-13)** only very modestly and positively correlates with awareness, while it seems not to have any effect on the probability for citizens to perceive either a personal benefit or appreciate EU actions through policies. This result points to the decaying effectiveness of communication in shaping the opinions of citizens over time.

Higher **proportions of negative vocabulary** (i.e. potentially carrying negative sentiments) in online communication significantly and negatively correlates with a lower probability of citizens appreciating the impact of EU policy interventions regarding both personal appreciation of the local policy implementation or membership of their countries in the EU (what we refer to as political appreciation). This result is highly consistent across several model specifications which include different combinations of topics discussed by local implementers through Facebook.

The amount of **negatively connoted topics** seems to generally lead to **less appreciation** in terms of both perceived personal benefit (i.e. see clusters of topics 3 and 6) and political appreciation (i.e. see cluster 2). Even though results are not always straightforward, the underlying methodology is quite innovative and holds potential in further developing our understanding of an eventual European public sphere.

0.3 Policy implications

A first and very general implication of our findings is that more attention should be paid from both policymakers and students of European integration to communication as a mechanism in amplifying awareness and especially appreciation of EU policies. Our analyses point to the fact that more funds translate into more awareness and appreciation.

A second implication directly follows by further disentangling time dynamic aspects of the first implication. That is, while the budget specifically allocated by LMAs for communication in the current period (2016 data) has an intuitive positive effect on personal appreciation for the actions of the EU, the funds allocated to communication tasks in the former programming period (2014 data) have no impact on appreciation and only very little impact on awareness. The implication for policy development and implementation here builds on the idea of a decay of the communication impact. That is, people in regions which have benefited of dedicated communication investment in the past, might experience a drop in their levels of awareness and appreciation after a decline in both structural funds and communication budgets (and efforts). It is suggested here that communication strategies specifically dealing with the objective of 'sustaining' awareness and appreciation levels in local territories should be employed once regions eventually advance to higher levels of development (GDP-based) and therefore experience declining levels of EU structural funding.

Finally, our research suggests that social media discourse might affect awareness and appreciation of EU policy in non-linear ways. Therefore, major attention should be devoted to developing

advanced and up-to date communicative skills at the local implementation level. As both citizens as well as policymakers seem to be embedded in collectively created structures of meaning (i.e. the post-comment interaction on social media), their attitude towards the EU might depend on this embedding. In light of this, we propose that communication ought to be more inspired by recent advancements on the strategic use of language in order to re-shape agency in fields and foster institutional change as well as the diffusion of new concepts and social identities.

0.4 Organization of the report

The document is organized as follows: The first section introduces the issue of assessing the effects of communication efforts in the context of EU public policy and political communication. Special attention is devoted to social media. The second section illustrates our empirical analytical strategy. The following third section reports our findings while the fifth and last section discusses the findings in light of relevant literature and policy implications.

1 State of the art

Awareness and appreciation by EU citizens

In previous deliverables (e.g. D5.4), we have focused on the degree to which individuals consider themselves Europeans and the meaning they attribute to their perception of being European (i.e. both European identification as well as the meaning of European identity). Project partners, too, have tested different factors leading to identification with the European Union (see deliverable 2.4, and deliverable 2.2).

Building on this prior work, we now steer the focus to indicators previously reviewed as independent variables leading to European identification, namely: a.) the awareness as well as b.) the appreciation of EU regional cohesion policy by EU citizens. As sketched out, we draw from both results of past deliverables as well as from literature on awareness and appreciation (in the context of EU policies) to describe the factors preceding what we consider dependent variables in this report.

Awareness of EU regional cohesion policy: Investment in structural funding and communication

As regards literature, we largely build on Osterloh (2011), who discusses variables affecting the level of awareness of citizens in the context of structural policy. Following an information-based approach (see also D2.2, 3.1. on mechanisms of European identity; or Bergbauer, 2018), the author emphasizes the amount of funding allocated to certain regions as creating awareness at the citizen level. More so, he describes the mechanism of political gains in the sense that legislators influence the distribution of funds to certain regions, regional citizens then become aware of the benefits created and reward the legislators for those benefits (p.4; hence, in our case, with implications for the appreciation of EU regional cohesion policy too). In this sense, this paper follows what we later refer to as **'rational choice hypothesis'**, in which citizens respond to economic incentives in the form of investments in regional and urban funding, instanced here by the **amount of structural funding**. In a similar manner, theories of economic utilitarianism (Fligstein, Polyakova, & Sandholtz, 2012; Verhaegen & Hooghe, 2015) assume that citizens will identify with the European Union following cost-benefit calculations. These can be considered in view of egocentric and sociotropic benefits (Hooghe & Marks, 2005), i.e. **benefits to oneself, as well as benefits to the region.** In an attempt to capture both, we consider the amount of total social funds pro capita.

Public awareness is however conditional upon certain factors such as visibility, and the attribution to the European Union as benefactor. Osterloh (2011) hence describes **the importance of communication** and the promotion of EU policies, for example through the use of media (p.10). In this regard, we consider the amount of money that goes into promoting the visibility of funds at the regional level, more specifically, the communication expenses pro capita.

Considering both economic benefits to the region and visibility as conditional for awareness, EU regional cohesion policy seems particularly suitable for our case. First, the amount of funding allocated to the regional level is both traceable and publicly available; second, the European Union is actively promoting the visibility of funds through communication and promotion activities.

We further draw from prior deliverables (D2.2, D2.4), in which Aiello et al. (Deliverable 2.4, 2018), too, describe the effect of the amount of structural funds expenses at regional level (in terms of per capita expenses)as directly affecting awareness of EU regional cohesion policy. López-Bazo & Royuela (Deliverable 2.2, 2017) further describe the importance of citizen's attentiveness to and interest in certain EU topics as well as their level of education as affecting awareness. With regard to Osterloh (2011), they point out that awareness is conditioned by socio-economic factors such as education. They further consider variables affecting the level of awareness, and consequently citizen's identification with Europe, to be: occupation, income, gender and age (p.60; and part of our control variables in the following). As regards their findings, López-Bazo & Royuela (2017) report that awareness is more frequent among respondents in less developed regions. They further describe that citizen's awareness is more likely when exposed to EU messages (and experiences).

Appreciation of EU regional cohesion policy: Semantics and topics

As has been pointed out in our conceptualization of awareness of EU regional cohesion policy, considerations of **economic utilitarianism** play both into awareness and appreciation at the citizen level. In this regard, we assume that funding allocated to a certain area will increase the likelihood of awareness (see also D2.4), and – in extension - also the likelihood of appreciating EU regional cohesion policy and the European Union.

However, appreciation will not just come about from the amount of money allocated to communication activities, but also from the form of communication - both in terms of sentiment and topics used. In a similar manner, Hofmann et al. (2013) emphasize that it is not merely the move onto social media (or more precisely, social networking sites), but the ways of communicating that will ultimately determine success. In this sense, we refer to our previous work in deliverable 5.4, in which we elaborated the impact of framing and citizen's attitude towards the European Union (e.g. p.9). In short, our assumption here is that **positive topics** of the EU and EU policies will create positive attitudes and appreciation at the citizen level (Vliegenthart et al., 2008; albeit in the previous deliverable used for media discourse, the same is true for the LMA's communication activities instanced by new social media).

This assumption stems from previous literature research (as regards communication in the context of public policy and political communication, please refer to deliverable 3.3 in which we provide a non-comprehensive literature review) acknowledging the practical importance of social media as a strategic communication channel (e.g. Ceccobelli & Siewert, 2016; Gausis, 2017; Usherwood & Wright, 2017), but saying little regarding how communication contributes to the social construction of European policy, integration, an identity. We understand this social construction of public policy to coincide with the Europeanization of national discourses, i.e. the emergence of a European public sphere. As put in Risse's words: "the ability to communicate meaningfully across borders depends crucially on the extent to which the same issues are debated at the same time with similar frames of reference or meaning structures" (Risse, 2009, p.150).

Current research on the European public sphere is summarized by De Vreese (2007) by pointing out three research strands: a.) a "utopian" strand focused on the necessity and prerequisite of a European public sphere; b.) an "elitist" strand emphasizing instances in which a European public sphere could be observed; and c.) a "realist" stream centered on indicators of Europeanization in different national public spheres. Following up on one of the avenues of further research proposed

by the author, we here wish to advance on the measurement of a communicative European space based "not only [on] whether issues are addressed simultaneously, but also how these are discussed" (De Vreese, 2007, p.13) in view of topics, and the framing of issues. In light of this elaboration, we consider the average use of topics in a given semantic cluster in the following (for more information, see 2.2 Analysis).

Similar considerations hold true for the **sentiment** of communication (again, see also D5.4, p.10) aimed at uncovering people's emotions and attitudes communicated through language. In this regard, sentiment analysis has been largely used to monitor election campaigns (Tumasjan et al., 2010) or improve policy-makers' knowledge about their target audience (Ceron et al., 2014, Flores, 2017). In an attempt to capture the tone of communication at the local level, we consider the count of negative vocabulary as proportion of positive vocabulary.

Further independent variables: European identity

Along with prior work on European identification and meanings of European identity, we are further interest in the integration of the results of D5.4 and the current report. Along these lines, we review the interplay with identification levels (whether European, national, or regional) and definitions of European identity (whether perceived in civic or cultural terms; for more information, please refer to D5.4).

2 Social media analysis

In order to improve our understanding of the effects of communication efforts of LMAs on citizens' awareness and appreciation of EU regional cohesion policy, we have created an empirical design which is described in this section of the report. In very general terms, we modelled citizens' awareness and appreciation of EU cohesion policy as stemming from economic and communication investment as well as, more experimentally, from the content of communication efforts. Accordingly, after a detailed description of our data sources, our empirical design is structured in two main parts: first, we analyzed the content of LMAs' Facebook pages to derive semantic-based independent variables. Second, we used these semantic variables along more standard variables tracing EU investment in order to estimate regression models and statistically test correlations among variables of interest (i.e. the impact of investment and communication content on citizens' awareness and appreciation of the policy).

2.1 Data

This report combines different data sources, some of which have been generated in previous PERCEIVE deliverables. Others have been generated in the scope of the tasks covered in the current report itself. As regards the former, the PERCEIVE survey with EU citizens described in deliverable 1.2 was of crucial importance as it provided us with data on *awareness* and *appreciation,* which we used to measure the dependent variables for our regression models, as well as data on European identification and individual level control variables (i.e. gender, age, education, income etc.).

Other data were gathered from the regional database detailed in PERCEIVE deliverable 2.1 and including secondary statistical data of interest for the project as a whole (i.e. NUTS2 level variables such as amount of structural funds per capita, GDP, unemployment etc.). The main explanatory variable we derived from this dataset is the amount of structural funds per capita.

We also used secondary data made available from the Directorate-General for regional policy (DG REGIO) through their portal at the EC website[1]. The main information we extracted from these data is the policy communication investment related to different operational programs generally available at the NUTS2 level. In order to get the correspondence between 2007-2013 and 2014-2020 "priority themes" and "Intervention field" codes we referred to the officially reported source at the same portal[2].

Moreover, from our direct experience developed through and documented in the previous deliverables, we have learned that besides their empirical relevance, social media accounts of LMAs provide one of the most standardized sources of data describing how LMAs communicate and interact with their local audiences. Therefore, we used their Facebook pages in order to measure the content of their public posts.

An extensive description of our social media datasets as well as the methods of collection has been provided in deliverable 5.2. Therefore, we only provide a short description here. In short, we have collected data for analysis from Facebook. More specifically, we have retrieved contents from the respective pages of the 8 LMAs in the PERCEIVE case study regions[3] (distributed over 6 countries):

- Italy: Regione Emilia Romagna. We downloaded all the posts and comments from the general Facebook profile of the Region, as there is no specific Facebook profile devoted to communication and management of European funds under Cohesion Policy.
- Italy: Regione Calabria POR. In this case we downloaded all the posts and comments from the specific profile, managed by Regione Calabria, which deals with communication and management of European funds under Cohesion Policy.
- Austria: Regionalmanagement Burgenland GmbH, which is the official Facebook Profile of the Local Managing Authority.
- Poland: Official Facebook channel for Warmińsko-Mazurskie region's communication of Cohesion Policy.
- Poland: Official Facebook channel for Dolnośląskie region's communication of Cohesion Policy.
- Romania: Agentia pentru Dezvoltare Regionala Sud-Est (Sud Est Regional Development Agency); a specific profile devoted to Cohesion Policy.
- Sweden: We downloaded posts and comments from the Facebook official profile of Tillväxtverket, the Local Managing Authority.
- Spain: We downloaded posts and comments from the Facebook official profile of Junta de Extremadura, the Local Managing Authority. This profile is a generalist one.

In total, this amounted to 24.218 posts and 18.720 comments.

---

[1] Data can be found here: http://ec.europa.eu/regional_policy/EN/policy/evaluations/data-for-research/.
[2] Available here:
http://ec.europa.eu/regional_policy/sources/docgener/evaluation/data/categorisation_2014_2020_mapping.xls.
[3] Essex as case study region of the United Kingdom does not handle a Facebook channel.

Table 1 – Facebook data summary

	Facebook profile	first post	posts	comments	page likes	approx. post /year	likes/ post	likes/ comment	comments/ post
Italy	Emilia-Romagna	27/08/2009	3.379	5.210	41.132	422	27,45	0,60	1,54
	Calabria	12/05/2016	428	339	5.569	428	16,10	0,30	0,79
Austria	Burgenland	14/06/2014	578	68	268	193	3,76	0,40	0,12
Poland	Warmińsko-mazurskie	14/06/2012	1.777	8.319	24.527	355	30,93	0,53	4,68
	Dolnośląskie	10/06/2011	831	117	2.620	139	2,21	0,33	0,14
Romania	Regional level: Agentia pentru Dezvoltare Regionala Sud-Est	10/08/2015	551	22	689	276	5,60	0,00	0,04
	National level: Ministerul Dezvoltarii Regionale, Administratiei Publice si Fondurilor Europene	10/10/2013	4.687	1.339	12.968	1.339	9,82	0,33	0,29
	National level: Ministerul Fondurilor Europene	07/02/2013	1.634	1.143	23.207	384	22,80	0,47	0,70
Sweden	Tillväxtverket	22/02/2012	540	380	2.879	108	10,04	0,87	0,70
Spain	Junta de Extremadura	20/01/2012	16.134	4.265	18.787	2.933	22,03	0,44	0,26

2.2 Analysis - Inquiring the existence of a European Public Sphere

In previous deliverables, we carried out an analysis of the Facebook profile of each LMA. In particular, we described the use of topic modelling in deliverable 5.3 – i.e. the process through which we elicited a 20 topics model describing the discourse carried out in each profile through posts and comments. We described the topics used and the labels we induced to describe them. Then, in deliverable 3.3, we provided a basic description of the Facebook presence of LMAs. In particular, we addressed: a) the general focus of the page (i.e. specifically for Cohesion Policy or not); b) the content of the three most "liked" posts and comments in terms of the topics composing them. This qualitative description was then complemented with an international comparison over four indicators: 1) posts per year, 2) likes per post, 3) likes per comment and 4) comments per post.

In this deliverable, in terms of description, we moved a step ahead, and considered all the elicited topics to analyze similarities and differences among countries. In particular, we wanted to test whether a European Public Sphere could be mapped based on the presence of similar topics (bags of words used in communication) in different national contexts. Two main analytical techniques were used towards this end: first, the clustering of topics in order to detect internationally shared meanings and second, sentiment analysis in order to characterize the clusters in terms of their positive/negative polarity and so have a confirmative (or dis-confirmative) factor for the interpretation of clusters' content.

Clustering of topics

To this end, we modelled the topic models from previous deliverables as semantic networks where the dots are topics and links among them indicate shared words among topics (see figure 1). In more detail, we followed these analytical steps: first, we considered the 15 most important words per topic – the words that mainly define the topic. We translated them into English by using an automatic translation tool (i.e. Google translate). As the topics are bags of words, this choice seemed appropriate because automatic translators are by now very accurate when dealing with individual words. Second, we created a list of all the words obtained, and we binarized the links between topics and the list of concepts (i.e. 1 if a certain word describes a topic, 0 if otherwise). Third, we constructed an overall semantic space as a topic x topic table, counting words shared between topics (shared cells, or links, as we just defined them). Fourth, we used an algorithm called CONCOR (with three levels of split) to generate partitions in the overall semantic space described above. While the algorithm was originally designed for detecting structural equivalent roles and positions in social networks (White, Boorman & Breiger, 1976), it has recently been used in a series of semantic networks applications (i.e. Mohr, 1994; Jancsary et al., 2017). Applied to our data, CONCOR is expected to find clusters of similar topics based on the fact that they connect in a similar way with each other.

Sentiment analysis of Facebook topics

We observed different indices in terms of posts and comments indicating different ways in which LMAs co-produce contents together with their external audiences. We further disentangled these

key differences in terms of 'sentiment' potentially attached to the co-produced contents. That is, we wanted to see if contents produced by the LMAs somehow differ from contents produced by their external audiences in terms of positive, neutral and negative words used in the content production.

This task has been performed through so-called sentiment analysis. The main objective of this explorative technique is to assess which opinions, or sentiment, (i.e. positive, neutral and negative) are more associated with the respective text. In this vein, we have analysed all topics distinguishing the use made of then in both posts (i.e. as a proxy for the voice of the LMA) and comments (i.e. as a proxy for the voice of the external audiences) on the Facebook pages of LMAs.

Sentiment analysis is based on *lexicons* – collections of words coded according to sentiment they potentially express. We used several such lexicons in order to strengthen the reliability of our interpretation of results. The VADER lexicon in particular was used for its focus on social media (Hutto & Gilbert, 2014) while Liu was used for its diffusion in social sciences as well as presence in the literature we reviewed (Hu & Liu, 2004). While the two abovementioned lexicons are widely used in social sciences, they only provide lists of positive and negative words. Therefore, with the main issue of wanting to include neutral words, we used two extra lexicons: Sentic Net 4 (Cambria et al., 2016) and Subjective Clues (Wilson, Wiebe, & Hoffmann, 2005).

We did not use the lexicons in bundles with the semantic/syntactic parsers they are usually accompanied with. Instead, we used them as mere codes for individual words. This choice makes sense in the context of our analysis as we wanted to understand the prevalent sentiment of topics elicited before (in deliverable 5.3), in which topics were constituted by bags of words (therefore missing the textual context needed for parsers to work).

In practice, we have computed the prevalent sentiment of a given topic as a count of the matches between full lists[4] of words constituting the topics and the words as part of the four different lexicons coded per sentiment (i.e. positive, neutral and negative). In order to facilitate the interpretation of results we also computed ratios of negative/positive and neutral/positive words as the total number of negative (or neutral) words divided by the sum of positive and negative (or neutral) words.

3. Regression analysis

In the second part of our empirical design we aimed at a statistical test of the association between communication efforts of LMAs on the one hand and awareness of policy and appreciation of the EU among citizens on the other hand.

3.1 Dependent variables

Awareness: Awareness of EU citizens is derived from PERCEIVE survey questions Q1_1 – Q1_4 asking if the respondent has heard about any of the four different keywords referring to EU structural investment: i.e. *cohesion policy, regional policy, structural funds and financed projects in the local*

---

[4] The topic models we have computed are made of 100 words.

region. It is first computed as the sum of such mentions for each respondent and therefore can vary between 0 and 4 per respondent. Then it is recoded as a binary variable assuming the value 0 when no label is mentioned and 1 when at least one of the labels is mentioned. This variable is also used as independent variable in the models inquiring about appreciation as dependent variable.

Political appreciation: This variable measures appreciation for the EU as a political identity. It is a binary variable derived from recoding question Q8 of the PERCEIVE SURVEY: "*In general do you think that (YOUR COUNTRY'S) EU membership is: 1) a good thing, 2) a bad thing, 3) neither good nor bad, 4) not sure*". It therefore takes the value 1 when the answer was "a good thing" and 0 if otherwise.

Perceived personal benefit: The perceived personal benefit is considered a measure of both awareness and appreciation. It is derived by the PERCEIVE survey question Q3: *"To your knowledge, have you ever benefited in your daily life from any financed project?"* This variable is binary, assuming value 0 when the answer is no and 1 when the answer is yes*.*

3.2 Independent variables

The following independent variables are arranged according to their thematic relevance: investment comprises structural funds expenses as well as communication expenses, both pro capita; semantic variables comprise sentiment and average use of topics in a given semantic cluster; European identity variables refer to both identification levels and definitions of being European, as well as controls for political participation and satisfaction with the regional economy. Further control variables are listed in the end.

Investment

Structural funds pro capita (SF_pc)*:* This variable constitutes a reliable indicator of investment in the region of interest even if not directly imputable to communication efforts. It is calculated as the total amount of structural funding in the 2007-13 programming period divided by the average population in the same period and clustered at the NUTS2 region level. It is a numeric value expressed in Euro.

Communication expenses pro capita (proCap_comm_expense)*:* This indicator is derived from the expenditures reported by the directorate general regional policy (see https://cohesiondata.ec.europa.eu/) using the codes 123 for information and communication activities in the current programming period, and 86 for the last programming period (2007-2013). Values are cumulative within the policy programming periods and refer to 2014 to account for the former period. However, as the categorization of regions varies across periods and so does the funding structure, we have also included a second variable counted at the end of 2016 to account for the current period as well (***and proCap_comm_expense 2016***).

Semantics

Simple sentiment: The sentiment variable used is a measure of the polarity of communication content in the Facebook accounts of LMAs. It is computed per each LMA (equivalent to NUTS2 region) as the total amount of words constituting negative vocabulary divided by the total amount of words constituting positive vocabulary. As has been mentioned, the VADER Lexicon has been used because of its special orientation towards social media content. This measure encompasses both posts and comments.

Average use of topics in a given semantic cluster (TPC_CL … _P): This variable traces the LMAs' semantic spectrum. In other words, it expresses the communicative behavior of LMAs on Facebook as participation in an eventually international discourse. It is computed as the average use (considering Facebook posts only) of all topics which a given LMA has placed in a given cluster of topics. This variable was measured and tested only for the four clusters of topics with the highest level of internationality (i.e. including topics from the LMAs of at least four different countries).

European Identity

Identification level (Q9_1, 2 and 3): The identification variable measures the level of identification of EU citizens with the region, the member state, and Europe. It is a numeric variable with a value range of 0 – 10.

Definitions of being European (Q10): Definitions of being European measure the level of agreement of respondent regarding five different statements about the meaning of being members of the EU: *1) the right for EU citizen to live and work in any other EU country; 2) having the common Euro currency; 3) the Christian religion; 4) having a common European flag; 5) sharing a common European history and culture.* It is a numeric variable in which each statement can be rated from 0 (zero agreement) to 10 (total agreement).

Political participation (Q7): The amount of political participation is a proxy to control for the level of engagement with European politics. It is derived from the question Q7 of the PERCEIVE survey: *"Have you voted in either of the last two European elections?"* and was used as a factor assuming the values 1 (once), 2 (both) or 99 (none).

Satisfaction with regional economy (Q17): Satisfaction with the regional economy is a control derived from the PERCEIVE survey question Q17: *"In general, how satisfied are you with the economic situation in your region?"* There are four possible answers: *1 = very satisfied, 2 = somewhat satisfied, 3 = somewhat unsatisfied, and 4 = very unsatisfied.*

Other controls

A series of additional standard controls at the individual level were also implemented which are: ***Gender*** (D1: 0 = male, 1 = female); **Education** (recoded D2: 1=Elementary and High school or less, 2= Graduation from High school, 3 = Graduation from college, 4=Post Graduate degree); ***Age*** (recoded D3: 1=18-29, 2=30-49, 3=50-64, 65+); ***Employment*** (D5.a: 1=work in public sector, 2=work in private

sector, 3=self-employed/small business/freelancer, 4=unemployed, 5=housewife/houseman, 6=pensioner/retired, 7=Pupil/student/trainee, 8=unemployed other) and *Individual total household net-income* (D7).

3.3 Regression models

As all of our dependent variables are binary (i.e. 1 or 0) we used standard logistic regression models for testing the expected associations. Logistic regression, also called a logit model, is used to model dichotomous outcome variables. In the logit model, the log odds of the outcome is modeled as a linear combination of the predictor variables (see Hosmer & Lemeshow, 2000 for a detailed explanation). We used the open source statistical software R in order to run the tests[5]. The selected function in R is the "glm" or generalized linear model, where the "binomial" family is selected for logit regression.

We performed three main groups of tests: the first one had awareness as a dependent variable and investments, EU identity and controls as independent variables. The second group of tests is the most detailed. Here we ran regression tests with either political appreciation or perceived personal benefit as dependent variables, and investment, EU identity and controls as independents. The effect of awareness on appreciation was tested in separate models not including identification variables. This is because in spite of a low correlation with individual components, in the aggregate the variable awareness was losing its significance when including all identity elements. Finally, we tested the effect of different semantically derived variables (i.e. sentiment and clusters) over both political appreciation and perceived personal benefit. It is worth mentioning that as this empirical design (including clusters of topics as explanatory variables) is quite new, there is no consolidated knowledge about the effects to be expected. Therefore, in order to minimize possible interaction which would be difficult to disentangle, these models only include the semantic variables and the individual controls.

4 Findings

4.1 Inquiring the existence of a European Public Sphere: clustering of topics

After having analyzed the Facebook profile of each LMA in our sample in the previous deliverables, we now move forward to analyze similarities among topics elicited in different countries to investigate the possible existence of a European Public Sphere. Figure 1 below illustrates the semantic connectivity among topics or, in other words, the semantic network formed by topics as nodes and shared words between topics as links. The color of the nodes reflects the CONCOR clusters that will be discussed below. At first glance, we can already see that some clusters are more connected than others (see the central position in the overall map) and some are more international (see labels) than others.

---

[5] See documentation here http://stat.ethz.ch/R-manual/R-devel/library/stats/html/glm.html

Figure 1 – Clusters of topics in the aggregated semantic space
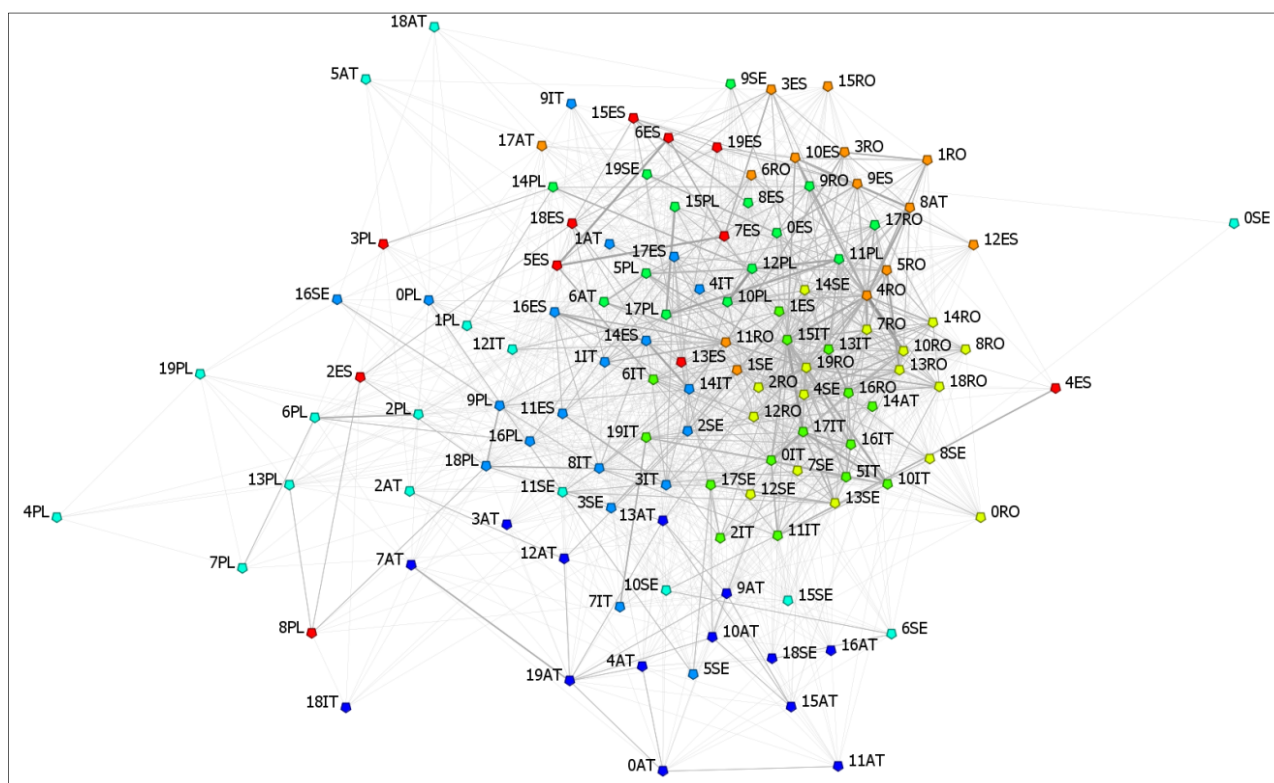
In the following 8 tables, we present the clusters of topics that we induced. The first column is a topic ID, the second column presents the name of each topic that was labeled together with project' partners. The third column presents the sentiment of the topic, which is calculated as a proportion of negative vocabulary used for that topic over positive vocabulary used for it. The fourth and fifth columns describe the usage of a certain topic for post and comments respectively. In example, the very first topic, which is topic 0 for the Austrian case, constitutes 5,2% of the post and 6,4% of the comments for the Austrian case.

Cluster 1: Austrian stylistic cluster

ID	NAME OF THE TOPIC	Sentiment	% post	% comment
0AT	COVERAGE OF PRESS ANNOUNCEMENTS	0.3	0,052	0,064
3AT	INFORMAL SETTING AT WORK	0.1	0,054	0,047
4AT	SPECIALIST EVENTS RESEARCH CLIMATE PROTECTION	0.1	0,052	0,047
7AT	BATH TOUR DURING SUMMER	0.2	0,052	0,046
9AT	COOPERATION TOGETHER	0.5	0,047	0,037
10AT	BEING PROUD, CONGRATULATIONS	0	0,057	0,042
11AT	THE PRESIDENT OF THE REGION	0.3	0,044	0,052
12AT	LMA WILL BE AT XYZ THIS YEAR	0	0,052	0,051
13AT	SOCIAL FUND AND COOPERATION	0.1	0,047	0,051
15AT	CROSS-BORDER PROJECTS	0	0,046	0,047
16AT	NOISE	0.1	0,042	0,039
19AT	TODAY, AGAIN, WE ARE DOING XYZ	0.2	0,064	0,042
18IT	NOISE	0.1	0,016	0,060
18SE	CONDITIONS FOR COMPANIES	0.5	0,053	0,046

This cluster collects 12 topics from the Austrian case, one from Italy, and one from Sweden. The reason for this clustering is probably self-explanatory, as it captures stylistic similarities in the use of vocabulary adopted in posts written by one Local Managing Authority, which is Regionalmanagement Burgenland GmbH. The usage of these topics for post and comments is quite balanced, with the only exception of 18IT, which indeed is a noisy topic. The only other non-Austrian topic deals with regional differences in growth and internationalization of companies.

Cluster 2: Euroscepticism

ID	NAME OF THE TOPIC	Sentiment	% post	% comments
1AT	ACTORS/SOCIAL RELATIONSHIPS	0	0,050	0,044
11ES	HEALTH SYSTEM AND SOCIAL SERVICES	1,5	0,05	0,041
14ES	DATA IN THE REGION OF EXTREMADURA	0.8	0,048	0,048
16ES	GENERAL COMPLAINTS	1	0,024	0,173
17ES	SOCIAL SERVICES	0.4	0,046	0,042
1IT	GENERAL COMPLAINTS	1.1	0,018	0,080
3IT	HEALTH CARE ADMINISTRATION	0.3	0,040	0,055
4IT	INFRASTRUCTURES	0.5	0,037	0,047
7IT	DISINFESTATION	0.7	0,021	0,062
8IT	SPECIFIC COMPLAINTS	0.5	0,018	0,089
9IT	VACCINE	1.8	0,019	0,064
14IT	YOUTH UNEMPLOYMENT	0.5	0,053	0,046
0PL	BRIDGE	0.1	0,042	0,048
9PL	ELBLAG CHANNEL	0.2	0,032	0,059
16PL	HOTEL	0.1	0,058	0,051
18PL	DIVERSE EVENTS	0	0,033	0,049
2SE	SOCIAL ENTERPRISES	0.4	0,059	0,046
3SE	INNOVATION AND ENTREPRENEURSHIP	0	0,045	0,047
5SE	INVESTMENT IN GLASSWORKS	0.4	0,029	0,070
16SE	QUESTION FOR THE PUBLIC ON INFRASTRUCTURE AND SAFETY	0.1	0,028	0,087

This cluster collects topics from 5 out of the 6 countries in our sample, being one of the less country-specific clusters. It is interesting to note that this cluster collects several topics related to citizens' complaints. . This is the case with Spanish topic 16 (16ES), which is about general complaints, exactly as 1IT, which collects laments by citizens generically directed at politicians. We also gather 8IT, which collects complaints targeted at specific episodes as, in example, the way Regione Emilia Romagna used European Funds for the reconstruction after the 2012 earthquake. These topics collecting complaints are more used in comments, than in posts. Interestingly, the topics that convey a critic are the most internationally distributed.

Together with complaints directed at European funds and their usage, this cluster collects topics characterized by negative sentiment. Topic 9IT, 'Vaccine' is characterized by the highest negative/positive vocabulary ratio. Indeed, this topic deals with a very heated debate in Italy, which followed the approval of a law to increase the number of mandatory vaccines for children. In this case, anti-vaccine followers were extensively using Facebook to support their stance. Apart from the clustering of complaints, topics around social services and healthcare seem prevalent: that is in the case of Spain (health system, social services), and Italy (health care administration, vaccine, disinfestations). Also social aspects in general are treated here.

Cluster 3: Tourism, culture, events

ID	NAME OF THE TOPIC	Sentiment	% post	% comment
2AT	LUDIC EVENTS	0.1	0,066	0,044
5AT	SUMMER FEST WITH MUSICIANS	0.2	0,046	0,051
18AT	POSSIBILITY TO WIN SOMETHING FROM LMAs	0	0,042	0,043
12IT	MANAGING STRUCTURAL FUNDS IN CALABRIA	0.6	0,056	0,054
1PL	OPEN DAYS OF EU FUNDS	0	0,066	0,039
2PL	HOTEL	0	0,042	0,050
4PL	MAZURIE FOR ALL SEASONS	0.1	0,030	0,062
6PL	ANPHITEATHER	0.1	0,030	0,062
7PL	REGIONAL AMENITIES	0.1	0,030	0,063
13PL	LOVE WARMIA E MAZURI	0	0,035	0,070
19PL	PHOTOS	0.1	0,039	0,061
0SE	MISUSE OF TAX MONEYS	1.6	0,023	0,096
6SE	TOURISM AND GROWTH	0	0,058	0,040
10SE	INTERNATIONALIZATION AND GROWTH	0	0,057	0,042
11SE	STUDENT ENTREPRENEURS	0	0,052	0,042
15SE	CONFERENCE OF THE BALTIC SEA REGION	0.4	0,039	0,047

This cluster collects topics from 4 countries. These topics are connected through the advertisement of events, festivals, cultural events, activities, and tourism. These topics are mainly used for posts by LMAs, with the exception of the Polish case. Here we find several topics, such as 4PL, 7PL and 13PL, which are used to describe the beauty of Mazury, of its dishes and of its amenities. These posts attract even more comments on the beauty of living in Warmia & Mazury. The only apparent intruder in this cluster is 0SE, which is the only topic with very negative sentiments, and which is mainly about Christina Lugnet, a Swedish politician that had to resign when it was discovered that she, in the role of Director General of Tillväxtverket, approved the expenditure of almost 7.5 million Swedish kronor (about 700.000€) for seminars and representation activities. The words used to describe these activities are probably similar to the ones describing events, used by the other topics of this cluster.

Cluster 4: Europe, funds, and programs

ID	NAME OF THE TOPIC	Sentiment	% post	% comment
6AT	COMMUNICATING EUROPE	0	0,043	0,085
0ES	EUROPEAN YOUTH POLICIY IN EXTREMADURA	0	0,055	0,033
8ES	GOVERNMENT AGREEMENTS	0	0,057	0,034
5PL	EU FUNDS	0.2	0,044	0,042
10PL	EU GRANTS	0.1	0,081	0,036
11PL	PROGRAMMES	0.3	0,086	0,036
12PL	JOB AND THE EU	0	0,076	0,036
14PL	FUNDS IN A LENS	0.1	0,086	0,042
15PL	FUNDS FOR INVESTMENTS	0	0,056	0,042
17PL	EU FUNDED RESCUE SERVICES	0.3	0,073	0,037
9RO	EU FUNDS	0	0,048	0,060
17RO	NOISE/COMPLAINTS	0	0,019	0,038
9SE	START UPS	0.1	0,052	0,047
19SE	GROWTH AND SOCIETY	0.2	0,044	0,049

This cluster, which is mainly about Communicating Europe, its funds and programs, collects topics from 5 countries, although Poland is the country bringing more topics to this cluster. Italy is the only country not represented here. Interestingly, topics in this cluster, on average, have the most positive sentiment in our sample.

AT6, which deals with events for communicators and Cohesion Policy implementers, is the only topic that is widely more used in comments, than in posts. Also, 9RO, which deals with instructions regarding clarifications for the technical and financial applications of funding through the Operational Program Human Capital, is more used in comments, although to a lesser extent. Conversely, all the other topics constitute posts, more than they constitute comments. Interestingly, the topics in this cluster, which generally aim at prospective beneficiaries (explaining EU funds, programs, funds in a lens, funds for investments), are rather composed of posts than comments. We expected more interaction, imagining people asking questions on funds, and programs. And yet, apparently, topics that inform on programs, and that are characterized by a positive sentiment, do not generate interactions with citizens.

Cluster 5: Funding, politics, and development

ID	NAME OF THE TOPIC	Sentiment	% post	% comment
8AT	ANNOUNCEMENTS	0.2	0,044	0,050
17AT	IMPRESSIONS FROM A RECENT EVENT	0.1	0,048	0,065
3ES	CONSTRUCTION SECTOR	0	0,057	0,035
9ES	BUSINESS SECTOR AND ENTREPRENEURSHIP	0	0,068	0,032
10ES	DEVELOPMENT PROGRAMMES	0	0,056	0,034
12ES	AGRICULTURE AND ENVIRONMENTAL ISSUES	0.4	0,054	0,037
1RO	ACCELERATING INFRASTRUCTURE WORKS TO OBTAIN FUNDING	0.1	0,033	0,187
3RO	EU FUNDING: ISSUES WITH FUNDINGS	0.2	0,017	0,046
4RO	EU FUNDING: PROCEDURES FOR SME	0	0,029	0,030
5RO	POLITICS: MEETING WITH CHINA	0	0,029	0,024
6RO	DISPUTE ON REGIONAL DEVELOPMENT AGENCIES' REMUNERATION	0.4	0,014	0,037
11RO	POLITICAL SCANDAL INVOLVING THE FORMER PRIME MINISTER	0.3	0,010	0,045
15RO	EU FUNDING PROCEDURES	0	0,268	0,046
1SE	SUSTAINABLE URBAN DEVELOPMENT	0	0,062	0,042

The most prominent country in this cluster is Romania, which accounts for 7 topics, followed by Spain, which accounts for 4 topics. Also Austria and Sweden are represented. The emphasis, as regards the Romanian topics, is on funding, politics and legislation. In particular, controversial topics are characterized by a more negative sentiment and by being used in comments, more than in posts. This is the case of 3RO, which deals with bureaucratic issues with funding, 6RO, which deals with disputes on a new law regarding LMAs' remuneration, and 11RO, which revolves around a political scandal.

The emphasis on financial and political aspects of Cohesion policy is coherent with the emphasis of Spanish topics on 'hard' sectors of Cohesion Policy, such as construction and agriculture. The Swedish topic also deals with development, especially with an urban one. These topics are used especially for posts and are characterized by a positive sentiment.

Cluster 6: Dissemination, communication, and Italy

ID	NAME OF THE TOPIC	Sentiment	% post	% comment
14AT	INITIATIVES WITH SCHOOLS	0	0,052	0,055
1ES	OPINION ON CULTURAL ACTIVITIES IN EXTREMADURA	0	0,050	0,035
0IT	REGIONAL POLITICS IN EMILIA-ROMAGNA	0	0,080	0,036
2IT	EUROPE AND THE EARTHQUAKE	2.3	0,056	0,042
5IT	EMILIA-ROMAGNA REGION	0	0,052	0,047

6IT	LOCAL CULTURAL POLICY	0	0,082	0,035
10IT	EMILIA-ROMAGNA AND THE EXPO	0	0,060	0,037
11IT	LOCAL EDUCATION POLICY	0.2	0,060	0,039
13IT	USING STRUCTURAL FUNDS IN CALABRIA	0.2	0,083	0,036
15IT	USING STRUCTURAL FUNDS IN EMILIA-ROMAGNA	0.3	0,091	0,040
16IT	PAOLA GAZZOLO	1.6	0,057	0,039
17IT	GENDER EQUALITY	0.3	0,075	0,038
19IT	REGIONAL VOTING	0.3	0,027	0,055
16RO	EU FUNDED PROJECT DISSEMINATION	0	0,175	0,064
17SE	EU FUNDED PROGRAMMES	1.3	0,058	0,046

This cluster is very Italy-centered, collecting 11 topics from this country, but further collects topics from Austria, Spain, Romania, and Sweden. Poland is the only country missing. Another trait of the topics pertaining to this cluster is their focus on dissemination and communication activities. Topic 14AT deals with initiative with schools, while the Spanish topic 1 provides information related to cultural activities carried out in Extremadura. Also, 16RO refers to a dissemination session regarding the implementation stage and the results of projects funded from European funds through the Regional Operational Program 2007-2013 at the level of the Sud-Est development region. 17SE, as well, is aimed at sharing information about the benefits of EU programs. Comparing the Italian topics with the other, we also find similar topics regarding education (14AT and 11IT) and cultural aspects (1ES, 6IT). As far as Italy is concerned, there is a clear focus on the two regions of Emilia-Romagna and Calabria, with several topics emerging that are related to the local dimensions. Generally speaking, the topics are used more for posts, than for comments. Negative sentiment emerges only in connection with the protection of Emilia Romagna: topic 2IT deals with the earthquake that hit the region in 2012, whereas topic 16IT deals with the activities of Paola Gazzolo, who is responsible for the protection of the territory and the coast of the Emilia Romagna region. Words such as "emergency", "bad weather", "security" and "territory" highlight this emphasis.

Cluster 7: Tourism, culture, and Spain

ID	NAME OF THE TOPIC	Sentiment	% post	% comment
2ES	NOISE	0.1	0,026	0,113
4ES	EDUCATION	0.3	0,055	0,037
5ES	CULTURAL ACTIVITIES: DATES AND PLACES	0.3	0,056	0,034
6ES	CULTURAL ACTIVITIES: PEOPLE	0.5	0,047	0,051
7ES	TOURISM AND EMERGENCY ALERTS	0.3	0,052	0,034
13ES	ECONOMICS AND SOCIETY	0.1	0,054	0,047
15ES	AGENDA OF THE PRESIDENT OF EXTREMADURA	0	0,054	0,036
18ES	LETTERS TO THE JUNTA	0.1	0,035	0,069
19ES	SECURITY AND CIVIL PROTECTION	0.1	0,055	0,035
3PL	CONGRATULATIONS	0.1	0,032	0,6
8PL	REGIONS TOURIST ATTRACTION	0	0,031	0,57

This cluster is very Spain-centered, while the only other country represented is Poland. Tourism and culture go together (7ES, 5ES, 6ES) and also economics and society (13ES) are mentioned. It is also important to denote the inclusion of both the Agenda of the president of Extremadura (15ES) and the letters to the Junta (18 ES). The latter is strangely only very little negative (0.1) as from the interpretation we read that comments from citizens are more complaints than praises. This topic is the only one, among the Spanish, which is widely more used in comments, than in posts. The association between 7ES and 8PL makes sense on the basis of tourism; this Polish post is more used for comments, than for posts.

Cluster 8: Technicalities on funding, entrepreneurship, and cultural projects

ID	NAME OF THE TOPIC	Sentiment	% post	% comment
0RO	EMERGENCY WARNING	2	0,029	0,027
2RO	NOISE: CELEBRATION OF ROMANIA'S NATIONAL DAY	0.1	0,019	0,043
7RO	INSTRUCTING CAMPAIGN	0.3	0,040	0,040
8RO	LOCAL PUBLIC INVESTMENTS	0	0,021	0,037
10RO	EU FUNDING	0	0,110	0,050
12RO	EU FUNDS	0.4	0,024	0,031
13RO	LEGISLATIVE PROCESS	0.1	0,039	0,044
14RO	INTEGRATED DEVELOPMENTS	0.8	0,035	0,042
18RO	CULTURAL HERITAGE	0.1	0,020	0,042
19RO	NOISE: EUROPEAN BLOGGING COMPETITION	0.1	0,020	0,067
4SE	MINING PROJECTS	0.2	0,054	0,054
7SE	START UPS	0.1	0,052	0,043
8SE	ENTREPENEURSHIP AND TEACHERS	0.1	0,057	0,039
12SE	CULTURAL AND CREATIVE COMPANIES AND EXPORTS	0.2	0,052	0,038
13SE	JOB ADS FOR STRUCTURAL FUNDS RELATED EMPLOYMENT	0	0,060	0,039
14SE	ENVIRONMENT AND POVERTY REDUCTION PROGRAMMES	0	0,066	0,039

This cluster collects topic from the Romanian and Swedish cases only. These topics are mostly characterized by the fact that they provide details on funds and on how to participate to bids. It is the case of 10RO, this relates to some examples of EU funded projects (REGIO 2007-20013/2014-2020) for education, culture and leisure; and the case of 12RO, which refers to a series of problems related to accessing European funds and public consultation regarding the Financing Manual for "Precedential support for the elaboration of Local Development Strategies". Also, this cluster has a component of entrepreneurship, which is mainly apparent in Swedish topics 7, 8, and 12. The topics in this cluster are mainly used for posts, with the exceptions of 2RO, and 19RO. The former deals with celebrating Romania's National Day, and collects positive reactions. The latter deals with a European blogging competition, and was mainly used in post precisely aimed at triggering reactions in comments. Overall, the sentiment is mostly positive, with the exception of 0RO, which actually collects posts linked to hydrological warnings for a number of bodies of water from Romania.

4.2 Sentiment analysis of Facebook topics

In this section we illustrate the results of the sentiment analysis associated with the topics in posts and comments appearing on the LMAs' Facebook pages. The table below summarizes, for each sentiment lexicon we experimented with, the average ration of negative vocabulary over positive vocabulary across all topics per country. We decided to use the values of the VADER lexicon as an independent variable in our models for two main reasons: a) it is specifically suited for the analysis of social media contents and b) it results in a variable which varies more than most of the other cases across countries.

	SENTICNET4	VADER	LIU	SENTIWORDNET3	SUBJCLUES
IT	0.3747	0.4679	0.4737	0.6868	0.3575
ES	0.2947	0.2393	0.2523	0.6216	0.2283
PL	0.2688	0.0722	0.1333	0.4135	0.1595
RO	0.2805	0.1639	0.2375	0.5519	0.2222
SE	0.3535	0.2349	0.3125	0.5957	0.3018
AT	0.2819	0.0985	0.1098	0.5668	0.1263

4.3 Regression tests

Awareness

As a first step in our analysis, we tested the determinants of awareness of European policies by European citizens. We began by testing a rational choice hypothesis, that is, we investigated how and whether citizens respond to economic incentives. This analysis suggests that the allocation of structural funds does produce awareness. This finding is, of course, to be controlled for factors such as level of education and age. We found that the link between allocation of structural funds and awareness is stronger for higher levels of education and for people in their fifties. This finding indicates, we propose, that the link between structural fund allocation and awareness of EU policy needs results as the outcome of adequate communication and education processes. Indeed, our study reports that increasing expenses in cohesion policy's communication leads to an increase in citizens' awareness. In addition, our analysis confirms that the effectiveness of communication investments weakens as time goes by since the effect of the communication investments made in the current programming period on awareness is stronger than the effect of the investments made in the last programming period. The reported effect of funds allocation and of communication investments on awareness holds in the case of high awareness as well. That is, structural funds and communication investments are connected not only to awareness of only one amongst EU cohesion policies, EU Regional policies or Structural funds, but they are positively connected to the awareness by citizens of more than one of these policies.

Our analysis suggests that not only economic factors produce awareness. Rather, the analysis detects a group of citizens whose awareness of EU policies seems to be connected to political and identification processes. As regards political involvement, awareness is more likely as the number of having voted in European elections increases. On the other hand, when identification with the European Union and with a common European history and culture scores high, citizens are more aware of European policies. Here again, however, the role of education, census and age is important.

Table 2 – Effect of policy-related investment and communication budgets on citizens' awareness of EU structural policy

	Model 1				Model 2				Model 3			
	Estimate	Std. Error	z_value		Estimate	Std.Error	z_value		Estimate	Std. Error	z_value	
(Intercept)	1.3173	0.0833	15.8240	***	-0.0633	0.2283	-0.2770		-1.1804	0.3581	-3.2960	***
Structural Funds pc	0.0025	0.0005	5.2520	***	0.0031	0.0005	6.2840	***	0.0029	0.0005	5.5470	***
Communication expense 2014 pc	0.0465	0.0260	1.7850	.	0.0547	0.0271	2.0200	*	0.0678	0.0300	2.2620	*
Communication expense2016 pc	0.1128	0.0397	2.8410	**	0.1167	0.0424	2.7530	**	0.1226	0.0457	2.6820	**
Vote in last EU election									0.7773	0.1446	5.3750	***
Vote in second-last EU election									0.6690	0.1163	5.7530	***
Vote in none of the above									-0.2414	0.2887	-0.8360	
Identity level - region									0.0338	0.0223	1.5190	
Identity level - country									0.0275	0.0239	1.1490	
Identity level - Europe									0.0719	0.0222	3.2400	**
EU definition - mobility									0.0313	0.0220	1.4230	
EU definition - Euro currency									0.0161	0.0169	0.9540	
EU definition - Christian religion									0.0168	0.0152	1.1060	
EU definition - Flag									0.0228	0.0185	1.2330	
EU definition - history and culture									-0.0502	0.0191	-2.6240	**
Satisfaction Ec. Regio.									-0.0621	0.0646	-0.9610	
Gender					-0.0456	0.0986	-0.463		-0.0580	0.1014	-0.5720	
Education – high school					0.5438	0.1165	4.669	***	0.4647	0.1202	3.8670	***
Education – college					1.2787	0.1558	8.210	***	1.1331	0.1601	7.0760	***
Education – post graduate					1.6722	0.2305	7.254	***	1.4755	0.2362	6.2480	***
Age 30-49					0.2372	0.1512	1.568		0.0945	0.1584	0.5970	
Age 50-64					0.4623	0.1715	2.695	**	0.3096	0.1801	1.7200	.
Age 65+					0.3556	0.2243	1.586		0.1323	0.2342	0.5650	
D4					-0.0006	0.0004	-1.433		-0.0006	0.0004	-1.5390	
Occupation – public sector					-0.0320	0.1530	-0.209		-0.0903	0.1564	-0.5770	
Occupation – private sector					0.2500	0.1979	1.263		0.2496	0.2024	1.2330	
Occupation – small business, free lancer					0.4376	0.2257	1.939	.	0.5022	0.2337	2.1490	*
Occupation – unemployed					-0.0496	0.2299	-0.216		-0.1535	0.2353	-0.6520	
Occupation – housewife/man					0.3270	0.2050	1.595		0.2778	0.2101	1.3220	
Occupation – student					0.6166	0.2966	2.079	*	0.5886	0.3041	1.9360	.
Occupation – unemployed other					0.5050	0.4242	1.190		0.5909	0.4393	1.3450	
Individual income2					0.2410	0.1259	1.914	.	0.2320	0.1287	1.8030	.
Individual income3					0.5932	0.1307	4.538	***	0.5806	0.1350	4.2990	***
Individual income99					-0.2773	0.1840	-1.507		-0.2516	0.1901	-1.3240	

Perceived personal benefit

Our variable *perceived personal benefit* mixes elements of awareness and appreciation to a certain extent (see description of variables in the previous section of this report). As the effects in our tests are pretty stable across increasingly detailed models, we interpret model 3, which is the most complete.

We found that the effect of both the general amount of structural funds pro capita and expenses specifically allocated to communication initiatives positively correlate with a higher probability of perceived benefit in the case study regions. This result is rather intuitive, as the underlying idea follows that more funds should lead to more financed projects and more communication to the knowledge and eventually perceived benefit for the projects.

Because the variables measuring the investment in communication in the two distinct programming periods were correlated (0.4), we tried to include either or both of them in different tests. As the test with both of them did not significantly differ from those with either, we kept both of them in the final estimation. What we see here is that the more recent investment (2016 is also closer to the time the PERCEIVE survey was conducted) is more significant than the one referring to 2014, signaling a potentially decaying effect for the effectiveness of communication.

A far as European identity is concerned, the results are mostly in line with our expectations. Political participation is not surprisingly positively and significantly associated with the probability of assuming a perceived personal benefit from the EU financed projects. However, and different from the test over awareness, only the voting in the last election has a significant impact. All three levels of identification – regional, member state and EU – are positively and somewhat significantly associated with the probability of perceiving a benefit from the financed projects. The strongest effect is observed for the member state level.

Considering the definition of being members of the EU we observe a complex but potentially meaningful pattern of effects. The value of mobility is positively associated with the perceived benefit, while the Christian religion is negatively associated. While these results are somehow intuitive a significant effect for the historic and cultural dimension of being members of the EU deserves a bit more of interpretation. This result is somewhat counterintuitive as we have previously observed (i.e. see deliverable 5.4) the historic-cultural element as being potentially more characteristic of Eurosceptic individual profiles (as it happens for the Christian religion). However, this might also open new lines of interpretation as for example pointing to a perceived benefit deriving from the valorization of the cultural heritage of regions.

The general level of satisfaction is positively and significantly associated with the perceived benefit. The negative coefficient of the effect in model 3 is explained by the inverse direction of the levels of the variable. That is, lower values of the variable correspond to higher satisfaction for the current economic situation of the region.

As regards the effect of awareness, (see model 4), we observe a positive and significant correlation with the probability of sampling a respondent whose perception of benefits from EU -financed projects is positive. This result is in line with our and standard expectations in public policy communication.

Some of the controls are also significantly pointing to the importance of education and individual income levels as well as to the potential dissatisfaction of small entrepreneurs, freelancers and the unemployed.

Table 3 – Effect of economic variables, European identity and awareness of policy on personal appreciation

	Model 1			Model 2			Model 3			Model 4		
	Estimate	Std.Error	z_value	Estimate	Std.Error	z_value	Estimate	Std.	z_value	Estimate	Std.	z_value
(Intercept)	-1.8496	0.0781	-23.6670***	-2.5706	0.1964	-13.0910***	-2.9520	0.2983	-9.8950 ***	-2.8861	0.2026	-14.2450 ***
Structural Funds pc	0.0045	0.0004	12.7920 ***	0.0053	0.0004	13.5170 ***	0.0047	0.0004	11.3390 ***	0.0047	0.0004	12.0490 ***
Communication expense 2014 pc	-0.0081	0.0232	-0.3510	-0.0023	0.0235	-0.0990	-0.0012	0.0250	-0.0470	-0.0172	0.0238	-0.7220
Communication expense 2016 pc	0.2151	0.0270	7.9770 ***	0.1488	0.0296	5.0280 ***	0.1416	0.0323	4.3910 ***	0.1140	0.0300	3.7930 ***
Vote in last EU election							-0.0246	0.1112	-0.2210			
Vote in second-last EU election							0.2706	0.0934	2.8960 **			
Vote in none of the above							0.1827	0.2411	0.7580			
Identity level - region							0.0401	0.0181	2.2120 *			
Identity level - country							0.0554	0.0201	2.7620 **			
Identity level - Europe							0.0420	0.0178	2.3520 *			
EU definition - mobility							0.0491	0.0190	2.5860 **			
EU definition - Euro currency							-0.0148	0.0128	-1.1560			
EU definition - Christian religion							-0.0496	0.0113	-4.3870 ***			
EU definition - Flag							0.0158	0.0139	1.1360			
EU definition - history and culture							0.0522	0.0150	3.4880 ***			
Satisfaction Ec. Regio.							-0.3362	0.0495	-6.7970 ***			
someMention										0.2986	0.0281	10.6090 ***
Gender (f)				-0.0268	0.0739	-0.3620	-0.0362	0.0766	-0.4730	0.0262	0.0752	0.3480
Education – high school				0.1576	0.1083	1.4550	0.1084	0.1120	0.9680	0.0600	0.1104	0.5430
Education – college				0.6220	0.1192	5.2180 ***	0.5866	0.1240	4.7290 ***	0.4509	0.1218	3.7010 ***
Education – post graduate				1.2850	0.1331	9.6530 ***	1.0595	0.1403	7.5530 ***	1.0907	0.1359	8.0230 ***
Age 30-49				0.0655	0.1150	0.5700	0.0006	0.1207	0.0050	-0.0482	0.1173	-0.4110
Age 50-64				0.1494	0.1285	1.1620	0.0505	0.1361	0.3710	0.0134	0.1312	0.1020
Age 65+				-0.1458	0.1768	-0.8250	-0.2646	0.1861	-1.4220	-0.2690	0.1795	-1.4980
Occupation – public sector				-0.1186	0.1089	-1.0890	-0.1218	0.1123	-1.0850	-0.1261	0.1108	-1.1380
Occupation – private sector				0.2001	0.1331	1.5030	0.2395	0.1369	1.7490 .	0.1619	0.1349	1.2010
Occupation – small business, free lancer				-0.5832	0.1678	-3.4750 ***	-0.4302	0.1737	-2.4770 *	-0.6301	0.1709	-3.6870 ***
Occupation – unemployed				-0.8599	0.2113	-4.0700 ***	-0.8408	0.2183	-3.8510 ***	-0.8769	0.2133	-4.1110 ***
Occupation – housewife/man				-0.1446	0.1520	-0.9510	-0.2521	0.1574	-1.6020	-0.1568	0.1539	-1.0180
Occupation – student				-0.2701	0.2198	-1.2290	-0.2421	0.2259	-1.0720	-0.3747	0.2222	-1.6860 .
Occupation – unemployed other				0.0138	0.2512	0.0550	-0.0994	0.2577	-0.3860	-0.0338	0.2559	-0.1320
Individual income2				0.2081	0.1032	2.0170 *	0.1931	0.1065	1.8120 .	0.1599	0.1049	1.5250
Individual income3				0.6632	0.0988	6.7120 ***	0.5115	0.1025	4.9910 ***	0.6071	0.1004	6.0490 ***
Individual income99				0.5629	0.1334	4.2190 ***	0.4327	0.1388	3.1160 **	0.5335	0.1357	3.9300 ***

Table 4 – Effect of semantic variables on perceived personal benefit

	Model 1			Model 2			Model 3			Model 4		
	Estimate	Std.Error	z_value	Estimate	Std.Error	z_value	Estimate	Std.Error	z_value	Estimate	Std.Error	z_value
(Intercept)	-0.2189	0.2052	-1.0670									
SimpleSentiment	-5.7593	0.3160	-18.2270***									
Average Use of topics cluster 2	2.5340	2.2388	1.1320									
(Intercept)				0.0612	0.1832	0.3340						
SimpleSentiment				-5.0930	0.3370	-15.1120***						
Average Use of topics cluster 3				-8.3409	1.4501	-5.7520 ***						
(Intercept)							-2.2224	0.2503	-8.8800 ***			
SimpleSentiment							-1.5584	0.4630	-3.3660 ***			
Average Use of topics cluster 4							26.4590	2.0977	12.6130***			
(Intercept)										0.1183	0.1839	0.6430
SimpleSentiment										-4.8997	0.3137	-15.6200***
Average Use of topics cluster 6										-8.7336	0.8496	-10.2790***
Gender (f)	-0.0826	0.0741	-1.1150	-0.0841	0.0741	-1.1360	-0.0811	0.0758	-1.0690	-0.0835	0.0757	-1.1030
Education – high school	0.2354	0.1088	2.1640 *	0.2535	0.1081	2.3460 *	0.1641	0.1106	1.4830	0.2353	0.1097	2.1450 *
Education – college	0.5965	0.1194	4.9950 ***	0.5995	0.1187	5.0520 ***	0.6637	0.1210	5.4870 ***	0.7308	0.1212	6.0270 ***
Education – post graduate	1.2674	0.1324	9.5760 ***	1.2817	0.1321	9.7010 ***	1.0831	0.1358	7.9750 ***	1.1885	0.1341	8.8640 ***
Age 30-49	0.1274	0.1146	1.1110	0.0980	0.1152	0.8510	0.1355	0.1173	1.1550	0.1638	0.1172	1.3980
Age 50-64	0.1658	0.1279	1.2960	0.1438	0.1284	1.1200	0.1773	0.1308	1.3550	0.2239	0.1308	1.7120 .
Age 65+	-0.0519	0.1741	-0.2980	-0.0549	0.1740	-0.3150	-0.0352	0.1786	-0.1970	0.0011	0.1789	0.0060
Occupation - work in private sector	-0.2535	0.1095	-2.3150 *	-0.2592	0.1091	-2.3750 *	-0.2065	0.1114	-1.8530 .	-0.1973	0.1115	-1.7700 .
Occupation - self-employed/small business/freelancer	0.2314	0.1353	1.7100 .	0.2492	0.1358	1.8340 .	0.2766	0.1377	2.0080 *	0.2080	0.1367	1.5220
Occupation - unemployed	0.0035	0.1759	0.0200	-0.0236	0.1781	-0.1330	-0.0935	0.1811	-0.5160	-0.0429	0.1781	-0.2410
Occupation - housewife/houseman	-0.4077	0.2187	-1.8640 .	-0.4336	0.2201	-1.9700 *	-0.4452	0.2251	-1.9780 *	-0.4033	0.2219	-1.8180 .
Occupation - pensioner/retired	-0.3510	0.1510	-2.3240 *	-0.3487	0.1503	-2.3200 *	-0.3174	0.1543	-2.0560 *	-0.3142	0.1542	-2.0370 *
Occupation - Pupil/student/trainee	-0.2718	0.2216	-1.2260	-0.3106	0.2215	-1.4020	-0.0960	0.2228	-0.4310	0.0163	0.2259	0.0720
Occupation - unemployed other	-0.1362	0.2425	-0.5620	-0.0741	0.2433	-0.3050	-0.1428	0.2491	-0.5730	-0.2905	0.2453	-1.1840
Individual income2	0.3160	0.1052	3.0040 **	0.3221	0.1056	3.0510 **	0.2778	0.1075	2.5840 **	0.2844	0.1072	2.6540 **
Individual income3	0.3775	0.0978	3.8610 ***	0.4161	0.0981	4.2390 ***	0.4009	0.1001	4.0040 ***	0.3677	0.0993	3.7020 ***
Individual income99	0.4212	0.1326	3.1770 **	0.4717	0.1323	3.5650 ***	0.2788	0.1364	2.0450 *	0.1538	0.1361	1.1300

In a separate set of models, we explored the effects of variables measuring the semantic structure and tone of communication generated through the Facebook accounts of LMAs and the perceived personal benefit of citizens in our case study regions. As mentioned before, this empirical design (including clusters of topics as explanatory variables) is quite new and there is no consolidated knowledge about the effects to be expected. Therefore, in order to minimize possible interaction which would be difficult to disentangle, these models only include the semantic variables and the individual controls.

With this premise we observe that the effect of the variable 'simple sentiment' is stable and remains negative and significant across the four distinct models (each of which tests the significance of a different cluster of topics plus individual controls).

Model 1 is centered on the effect of cluster 2, one of the clusters with the highest proportion of negative vocabulary among those that we observed. The effect of using topics of Cluster 2 labelled 'Euroscepticism' in the posts of LMAs is not significant.

Model 2 is built to test the effect of using topics of cluster 3 labelled 'tourism, culture and events' is negative and significant. Because the overall sentiment of the cluster 3 is positive, this result might seem counterintuitive, however, when we consider the international composition of the cluster things change. In fact the topics of both Norra Mellansverige (SE) and Calabria (IT) contributing to the same cluster are highly negative and could explain this result.

Model 3 tests the effect of using topics constituting cluster 4 which as whole is about communicating Europe, programs and potential beneficiaries. This effect is strongly positive and significant. In this case, none of the topic members of this cluster are to be considered negative and all in all this evidence seems to point to the strategic importance of communicating the related topics.

Model 4 tests the effect of using topics that belong to cluster 6 with a strong participation of Italian topics and labelled 'dissemination and communication' is negative and significant. As in the case of cluster 3, the label of the topic might be deceptive in that the cluster indeed includes very negatively connoted topics. However, two of the most negative ones are to be attributed to non-regional policy relevant discourse (as they deal with emergency management relative to an earthquake in Emilia Romagna.

Political appreciation of country membership in the EU

As in the former battery of tests, we tested the effect of economic investment factors, identity and personal controls, but this time the dependent variable has been set to be what we referr to as *political appreciation*. As described in the methods section, this variable is "1" when respondents of the PERCEIVE survey agreed that membership of their country in the EU is a good thing (question Q8) while it takes the value "0" when they agreed to the opposite or were undecided.

The economic effect of investments seems to have some impact on citizens' appreciation of EU membership. More specifically, allocation of structural funds is positively associated with higher appreciation of EU membership. This effect goes into the same direction as the one in the models with personal benefit as dependent variable. The effect of communication on appreciation instead goes into the opposite direction and seems somehow counterintuitive. Our analysis suggests that

citizen appreciation increases as the expenses for communication of the past programming period increases. However, this effect is feeble and not significant. Communication expenses of the current programming period seem to have a negative effect on appreciation. In general, the effect of communication expenses on appreciation seems to be both unclear and unstable when other determinants are included in the model. This surprising result may point to different modes of effectives of communication strategies in different programming periods and to the overall adequacy of the content and means of communication. Moreover, it calls for the assessment of changes in the design and administration of communication strategies in the passing from programming period 2007-2013 to 2014-2020. Also, it has to be noticed that the composition of budgets for regions has changed over the two periods (i.e. Burgenland (AT) did not receive a NUTS2-operative plan in the current period as the LMA has been centralized to the national level to respond to decreasing levels of funding)

Interestingly, however, appreciation of EU policies seems to be less clearly correlated with communication investments than identification processes. The analysis clearly signals how appreciation might follow from a process of identification. The more citizens identify with Europe, the more they appreciate EU policies. More specifically, the identification based on a common flag and passport, a common history and culture, a common currency and the right to work and live in European countries correlates with the increase of appreciation. Among the identification mechanisms that we tested, only the identification with a common Christian religion does not relate to appreciation.

As far as awareness is concerned, we observe a positive significant effect on political appreciation. While this result is not surprising, it is interesting to notice that it goes into the same direction as the models on perceived personal benefit.

Table 5 – Effect of economic variables, EU Identity and awareness of policy on political appreciation

	Model 1			Model 2			Model 3			Model 4		
	Estimate	Std.Error	z_value	Estimate	Std. Error	z_value	Estimate	Std. Error	z_value	Estimate	Std.Error	z_value
(Intercept)	0.0207	0.0626	0.3310	-0.6473	0.1680	-3.8540 ***	-2.7683	0.2934	-9.4340 ***	-0.7319	0.1695	-4.3180 ***
Structural Funds pc	0.0037	0.0003	10.6730 ***	0.0041	0.0004	11.2640 ***	0.0039	0.0004	8.8260 ***	0.0039	0.0004	10.6880 ***
Communication expense 2014 pc	0.0055	0.0187	0.2940	0.0038	0.0193	0.1990	-0.0326	0.0238	-1.3680	-0.0028	0.0194	-0.1430
Communication expense 2016 pc	-0.1405	0.0259	-5.4250 ***	-0.1718	0.0276	-6.2140 ***	-0.1414	0.0340	-4.1620 ***	-0.1870	0.0280	-6.6790 ***
Vote in last EU election							0.1549	0.1075	1.4410			
Vote in second-last EU election							0.3042	0.0932	3.2630 **			
Vote in none of the above							0.3118	0.2565	1.2160			
Identity level - region							0.0081	0.0180	0.4520			
Identity level - country							0.0449	0.0196	2.2850 *			
Identity level - Europe							0.2021	0.0175	11.5750 ***			
EU definition - mobility							0.1664	0.0183	9.0720 ***			
EU definition - Euro currency							0.1595	0.0129	12.3770 ***			
EU definition - Christian religion							-0.0119	0.0118	-1.0120			
EU definition - Flag							0.0377	0.0142	2.6570 **			
EU definition - history and culture							-0.0480	0.0149	-3.2120 **			
Satisfaction Ec. Regio.							-0.4714	0.0502	-9.3880 ***			
someMention										0.1010	0.0246	4.1000 ***
Gender (f)				0.1481	0.0672	2.2040 *	0.1056	0.0771	1.3700	0.1639	0.0674	2.4320 *
Education – high school				0.1761	0.0898	1.9600 .	0.1443	0.1039	1.3890	0.1425	0.0905	1.5750
Education – college				0.3848	0.1042	3.6920 ***	0.3155	0.1206	2.6170 **	0.3255	0.1054	3.0870 **
Education – post graduate				0.7978	0.1277	6.2480 ***	0.6060	0.1477	4.1020 ***	0.7219	0.1292	5.5870 ***
Age 30-49				0.0991	0.1055	0.9390	-0.1108	0.1216	-0.9110	0.0651	0.1060	0.6150
Age 50-64				0.1930	0.1182	1.6320	-0.1385	0.1378	-1.0050	0.1509	0.1188	1.2700
Age 65+				0.0547	0.1569	0.3490	-0.2223	0.1823	-1.2190	0.0178	0.1574	0.1130
D4				-0.0012	0.0003	-3.6210 ***	-0.0013	0.0004	-3.5430 ***	-0.0012	0.0003	-3.6720 ***
Occupation - work in private sector				-0.0193	0.1045	-0.1840	-0.1003	0.1185	-0.8460	-0.0180	0.1047	-0.1720
Occupation - self-employed/small business/freelancer				-0.1348	0.1279	-1.0540	-0.1614	0.1469	-1.0990	-0.1529	0.1281	-1.1930
Occupation - unemployed				-0.4128	0.1472	-2.8050 **	-0.3549	0.1696	-2.0930 *	-0.4208	0.1477	-2.8490 **
Occupation - housewife/houseman				-0.1474	0.1698	-0.8680	-0.2022	0.1933	-1.0460	-0.1479	0.1702	-0.8690
Occupation - pensioner/retired				0.2430	0.1396	1.7410 .	0.0871	0.1593	0.5470	0.2408	0.1398	1.7220 .
Occupation - Pupil/student/trainee				0.2434	0.1999	1.2180	-0.1974	0.2279	-0.8660	0.2142	0.2004	1.0690
Occupation - unemployed other				0.0799	0.2421	0.3300	0.1316	0.2787	0.4720	0.0637	0.2428	0.2620
Individual income2				0.0671	0.0885	0.7590	0.0409	0.1011	0.4050	0.0462	0.0888	0.5200
Individual income3				0.5495	0.0875	6.2830 ***	0.3470	0.1008	3.4420 ***	0.5213	0.0878	5.9340 ***
Individual income99				0.4447	0.1293	3.4380 ***	0.3535	0.1490	2.3730 *	0.4306	0.1297	3.3200 ***

Table 6 – Effect of semantic variables on political appreciation

	model 1			model 2			model 3			model 4		
	Estimate	Std.Error	z_value	Estimate	Std.Error	z_value	Estimate	Std.Error	z_value	Estimate	Std.Error	z_value
(Intercept)	1.2524	0.1907	6.5680 ***									
SimpleSentiment	-4.2453	0.2376	-17.8700***									
Average Use of topics cluster 2	-10.1361	2.3280	-4.3540 ***									
(Intercept)				0.8683	0.1630	5.3290 ***						
SimpleSentiment				-3.1837	0.2673	-11.9090***						
Average Use of topics cluster 3				-5.6362	0.9445	-5.9670 ***						
(Intercept)							0.8182	0.2363	3.4630 ***			
SimpleSentiment							-4.0054	0.4200	-9.5360 ***			
Average Use of topics cluster 4							0.0892	2.1056	0.0420			
(Intercept)										0.6670	0.1658	4.0220 ***
SimpleSentiment										-4.1854	0.2338	-17.901 ***
Average Use of topics cluster 6										3.5271	0.7642	4.6150 ***
Gender (f)	0.1163	0.0687	1.6940 .	0.1173	0.0687	1.7060 .	0.1187	0.0685	1.7330 .	0.1161	0.0687	1.6900 .
Education – high school	0.0759	0.0916	0.8290	0.1684	0.0915	1.8400 .	0.1220	0.0909	1.3410	0.1113	0.0911	1.2210
Education – college	0.3206	0.1063	3.0150 **	0.3930	0.1063	3.6970 ***	0.3608	0.1061	3.4020 ***	0.3172	0.1067	2.9730 **
Education – post graduate	0.4797	0.1295	3.7050 ***	0.5867	0.1294	4.5350 ***	0.5291	0.1297	4.0800 ***	0.5745	0.1291	4.4500 ***
Age 30-49	0.2073	0.1078	1.9240 .	0.1947	0.1080	1.8030 .	0.2066	0.1076	1.9200 .	0.2008	0.1076	1.8660 .
Age 50-64	0.2477	0.1198	2.0690 *	0.2438	0.1200	2.0310 *	0.2494	0.1196	2.0860 *	0.2412	0.1197	2.0140 *
Age 65+	0.2443	0.1604	1.5230	0.2434	0.1603	1.5180	0.2339	0.1599	1.4620	0.2267	0.1604	1.4140
Occupation - work in private sector	-0.1223	0.1068	-1.1450	-0.1071	0.1068	-1.0030	-0.0949	0.1065	-0.8900	-0.1158	0.1068	-1.0850
Occupation - self-employed/small business/freelancer	-0.1616	0.1304	-1.2390	-0.1607	0.1307	-1.2300	-0.1572	0.1304	-1.2060	-0.1460	0.1305	-1.1190
Occupation - unemployed	-0.0494	0.1508	-0.3280	-0.0009	0.1533	-0.0060	-0.0230	0.1507	-0.1520	-0.0173	0.1508	-0.1150
Occupation - housewife/houseman	0.0802	0.1736	0.4620	0.1563	0.1760	0.8880	0.1248	0.1730	0.7210	0.1165	0.1737	0.6710
Occupation - pensioner/retired	-0.0230	0.1426	-0.1610	0.0216	0.1422	0.1520	0.0272	0.1420	0.1920	-0.0046	0.1423	-0.0320
Occupation - Pupil/student/trainee	0.3324	0.2049	1.6220	0.4250	0.2060	2.0630 *	0.4000	0.2038	1.9630 *	0.3127	0.2062	1.5160
Occupation - unemployed other	-0.2987	0.2429	-1.2300	-0.2836	0.2417	-1.1730	-0.3211	0.2423	-1.3250	-0.2480	0.2423	-1.0240
Individual income2	0.1518	0.0905	1.6770 .	0.1435	0.0912	1.5730	0.1353	0.0904	1.4970	0.1591	0.0908	1.7520 .
Individual income3	0.4066	0.0884	4.5980 ***	0.3839	0.0884	4.3420 ***	0.3858	0.0883	4.3720 ***	0.4067	0.0884	4.5980 ***
Individual income99	0.0760	0.1302	0.5840	0.0675	0.1299	0.5200	0.0312	0.1304	0.2400	0.1652	0.1322	1.2500

Beside economic factors and identification, we tested the connection between the appreciation and the sentiment and structure of the discourse that occurs in the Facebook page of LMAs in the specific regions of each respondent. As expected, a general result is that a decrease in appreciation follows from detecting a negative sentiment in the discourse that takes place in the Facebook pages. Furthermore, we found a strong and significant negative relationship between appreciation and the presence of internationally diffused Eurosceptic discourse in LMAs' posts. Thus, negative sentiment is more effective in reducing appreciation the more it is conveyed through a discourse which is diffused in different European countries. The relationship between appreciation and discourse works both towards, generally, the appreciation of being part of the EU and towards, specifically, the appreciation of the benefit that accrued to the specific respondent's region.

We tested whether the opposite was true. That is, we tested whether the presence of internationally diffused pro-Euro discourse in the Facebook pages of LMAs leads to more appreciation. Interestingly, we discovered that an LMA's connection to an internationally positive discourse is a predictor of general appreciation of a country belonging to the EU and of the benefits of specific respondents' regions.

5 Discussion and policy implications

The PERCEIVE project investigates awareness and appreciation of EU policies adopting a social constructivist perspective. We focused on the discursive construction of the European integration process and European identities (Checkel, 1999, 2001, 2005; Diez, 1999; Paasi, 2001). Specifically, the project takes issue with the mainstream rational choice perspective in the literature on EU integration that stresses the idea of institutions as 'rules of the game' and the calculative rationality of actors. Following this perspective, citizens' decision to identify with the EU, for example, is regarded as the outcome of an evaluation cost-benefit functions associated with different systems of rules – i.e. frameworks for cohesion – and their observed outcomes – i.e. the implementation in beneficiary regions.

Our findings accommodate a social constructivist and rational-choice view. Specifically, we suggest that a rational choice perspective explains the positive effect that the allocation of structural funds to specific European regions has on citizens' appreciation and awareness. According to our findings, the allocation of structural funds produces awareness and appreciation of both membership to the EU and the benefits accruing to regions.

However, our findings unveil the role of identification processes as well as appreciation and awareness seem to be strongly correlated therewith. Citizens who identify with the EU are aware of EU policies and appreciate both EU membership and the benefits that accrue to European regions.

Furthermore, an important distinction that we captured concerns the type of EU identity that is activated by citizens. Known types of EU identity are civic, cultural and ethnic (Bruter, 2003, 2005; Fligstein, Polyakova & Sandholtz, 2012; for more information, see also deliverable 5.4). Scholars define civic EU identity as the degree to which people feel that they are citizens of a European political system, whose rules, laws, and rights have an influence on their daily life. To put it yet another way, civic identity refers to citizens' identification with their political system as an institutional frame, that is, their state.

Our analysis highlights that such a civic identity explains awareness of EU policies and appreciation of EU membership. In particular, we found a positive correlation between citizens' identification with the Euro, the freedom to live and work in European countries, on the one hand, and awareness and appreciation, on the other. We suggest that the Euro and the freedom to live and work in European countries represent the EU political system in which citizens may identify and well capture the essence of a civic identity. In addition, to confirm that a civic identity is at work, we found that awareness and appreciation is positively correlated with political activity; the latter measured as repeatedly voting in European elections.

Our research elicited a role for cultural identity as well. This is best described as individuals' perceptions of proximity towards fellow Europeans rather than non-Europeans. Thus, cultural identity refers to citizens' identification with their political community as a human group, regardless of the nature of the political system. Similarly, when talking about the possibility of cultural identification of EU citizens with Europe, we address the repertoire of common features that can be identified by individuals as creating a bond among Europeans. The perception of shared European heritage might include any form of common history; moral, religious, or ethnic traditions; philosophical, political, or moral norms and values; and so forth. Specifically, we found that while identifying with a common history is positively connected to awareness and appreciation, identifying with a common Christian religion does not correspond with awareness and appreciation. These findings suggest that cultural identity is a complex construct with different dimensions; the latter not necessarily working in the same direction. We suggest that these findings open the way for further research aimed at exploring the multifaceted composition of cultural identity.

As regards the communication analysis, our research suggests that a simple link between communication investment, and awareness and appreciation does not exist. Rather, citizens seem to react to more complex features of discourse. Having tried to address this issue, we analyzed the threads of discussion of LMAs' Facebook pages. First, we found that the prevalence of negative sentiment in a LMA's Facebook page is negatively connected to awareness and appreciation. Second, we found the existence of discourses that are articulated in different topics and distributed in different countries. These discourses may be Eurosceptic or Euro-optimistic in nature. More importantly, in those regions in which the Facebook pages of LMAs connect to Euro-optimistic discourses, we find that citizens are more aware of EU policies and appreciate EU membership. On the other hand, the presence of Eurosceptic discourses in the LMA's Facebook page anticipates a lack of awareness and appreciation by European citizens. What emerges from our research is that citizens as well as policymakers seem to be embedded in a structure of meaning that influences their attitude towards the EU. In light of these findings, we propose that communication analysis ought to be inspired by recent advancements on the strategic use of language in order to re-shape agency in fields and foster institutional change as well as the diffusion of new concepts and social identities (Cornelissen et. al., 2015; Harmon, Green & Goodnight, 2015; Suddaby & Greenwood, 2005; Vaara, 2014).

Therefore, in addressing the mechanisms of institutionalization of EU policies, our work resonates with the suggestion of Cornelissen & Werner (2014), who encourage scholars to study discursive opportunity structures (Koopmans & Statham, 1999, p.231; McCammon et al., 2007, p.745) as "the opportunity provided by salient discourses that are alive and have momentum at a particular point in time" (2014, p.210).

Taking this perspective, the analysis of communication, we suggest, requires appropriate skills to recombine 'available discursive elements' (Hensmans, 2003, p.362) and navigate the cultural texture of logics of a field, that is, its deeply ingrained societal beliefs and practices (Friedland & Alford, 1991). The 'latent meaning structure' of a field (Meyer & Höllerer, 2010), the 'multiplicity' of the field's logics (Hoffman, 1999) and the ideological incompatibility (Rao & Kenney, 2008) and 'relative incoherence' of the latter influence how citizens interpret the role of the EU and EU policies.

In conclusion, the social constructivist perspective of PERCEIVE stresses the idea of institutions as "constitutive of actors and preferences" and the social rationality of actors. Following this perspective, citizens' identification with the EU, for example, is regarded as the outcome of a social learning process associated with different institutional discourses and underlying logics – i.e. frameworks for cohesion – and their discursively constructed outcomes – i.e. the meanings associated with implementing projects in beneficiary regions.

6 References

Aiello, V., Brasili, C., & Calia, P. (2018). *Report on the probabilistic model of estimation of citizens' identification with the EU project and ranking of the case study regions.* PERCEIVE Deliverable 2.4. July 2018.

Bergbauer, S. (2018). *Explaining European identity formation: Citizens' attachment from Maastricht Treaty to crisis.* Springer.

Bruter, M. (2003). Winning hearts and minds for Europe: The impact of news and symbols on civic and cultural European identity. *Comparative Political Studies, 36*(10), 1148-1179.

Bruter, M. (2005). *Citizens of Europe? The emergence of a mass European identity.* New York: Palgrave Macmillan.

Checkel, J. (1999). Social construction and integration. *Journal of European Public Policy, 6*(4), 545-560.

Cambria, E., Poria, S., Bajpai, R., & Schuller, B. W. (2016). SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives. *COLING,* 2666-2677.

Ceccobelli, D., & Siewert, M. B. (2016). Campaigning on Facebook: How do candidates running for the presidency of the European Commission behave? *Conference Paper.*

Ceron, A., Curini, L., & Iacus, S.M. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New media & society, 16*(2), 340-358.

Checkel, J. (2001). Why comply? Social learning and European identity change. *International Organization, 55*(3), 553-588.

Checkel, J. (2005). International institutions and socialization in Europe: Introduction and framework. *International Organization, 59*(4), 801-826.

Cornelissen, J. P., & Werner, M. D. (2014). Putting framing in perspective: A review of framing and frame analysis across the management and organizational literature. *The Academy of Management Annals, 8*, 181–235.

Cornelissen, J. P., Durand, R., Fiss, P. C., Lammers, J. C., & Vaara, E. (2015). Putting communication front and center in institutional theory and analysis. *Academy of Management Review, 40*(1): Introduction.

De Vreese, C.H. (2007). The EU as a public sphere. *Living Reviews in European Governance*, 2(3).

Diez, T. (1999). Speaking 'Europe': the politics of integration discourse. *Journal of European Public Policy, 6*(4), 598-613.

Fligstein, N., Polyakova, A., & Sandholtz, W. (2012). European integration, nationalism and European identity. *Journal of Common Market Studies, 50,* 106-122.

Flores, R.D. (2017). Do anti-immigrant laws shape public sentiment? A study of Arizona's SB 1070 using twitter data. *American Journal of Sociology, 123*(2), 333-384.

Friedland, R., & Alford, R.R. (1991). Bringing society back in: Symbols, practices, and institutional contradictions. In W.W. Powell & P.J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 232–266). Chicago: University of Chicago Press.

Gausis, E. (2017). European institutions on social media: Shaping the notion of European citizenship. *Economics and Business, 30,* 27-39.

Harmon, D. J., Green, S. E., & Goodnight, G. T. (2015). A model of rhetorical legitimation: The structure of communication and cognition underlying institutional maintenance and change. *Academy of Management Review, 40*(1): 76-95.

Hensmans, M. (2003). Social Movement Organizations: A Metaphor for Strategic Actors in Institutional Fields. *Organization Studies, 24,* 355–381.

Hofmann, S., Beverungen, D., Räckers, M., & Becker, J. (2013). What makes local governments' online communication successful? Insights from a multi-method analysis of Facebook. *Government Information Quarterly, 30*, 387-396.

Hoffman, A.J. (1999). Institutional evolution and change: Environmentalism and the U.S. chemical industry. *Academy of Management Journal, 42*(4): 351-371.

Hooghe, L., & Marks, G. (2001). *Multilevel governance and European integration.* New York: Rowman & Littlefield.

Hosmer, D. & Lemeshow, S. (2000). *Applied Logistic Regression (Second Edition).* New York: John Wiley & Sons, Inc.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 168-177.

Hutto, C.J., & Gilbert, E.E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.*

Jancsary, D., Meyer, R. E., Höllerer, M. A., & Barberio, V. (2017). Toward a structural model of organizational-level institutional pluralism and logic interconnectedness. *Organization Science*, *28*(6), 1150-1167.

Koopmans, R., and Statham, P. (1999). Ethnic and civic conceptions of nationhood and the differential success of the extreme right in Germany and Italy. In M. Giugni, D. McAdam, and C. Tilly (Eds.), *How social movements matter* (pp. 225–251). University of Minnesota Press: Minneapolis, MN.

López-Bazo, E., & Royuela, V. (2017). *Mapping the determinants of EU citizen's perception and identification.* PERCEIVE Deliverable 2.2. December 2017.

McCammon, H., J.,Newman, H. D., Sanders Muse, C., & Terrell, T.M. (2007). Movement framing and discursive opportunity structures: The political successes of the u.s. women's jury movements. *American Sociological Review, 72*(5): 725–749

Meyer, R. E., & Höllerer, M. A. (2010). Meaning structures in a contested issue field: A topographic map of shareholder value in Austria. *Academy of Management Journal, 53*(6), 1241-1262.

Mohr, J. W. (1994). Soldiers, mothers, tramps and others: Discourse roles in the 1907 New York City charity directory. *Poetics, 22*(4), 327-357.

Osterloh, S. (2011). *Can regional transfers buy public support? Evidence from structural policy.* ZEW Discussion Papers 11-011, ZEW-Zentrum für Europäische Wirtschaftsforschung.

Paasi, A. (2001). Europe as a social process and discourse: Considerations of place, boundaries and identity. European Urban and Regional Studies, 8(1), 8-28.

Rao, H., & Kenney, M. (2008). New forms as settlements. In R. Greenwood, C. Oliver, K. Sahlin, and R. Suddaby (Eds.), *The SAGE Handbook of Organizational Institutionalism* (pp. 651–672). Sage: London.

Risse, T. (2009). Social constructivism and European integration. In A. Wiener & T. Diez (Eds), *European integration theory* (2nd edition, pp. 144-159). Oxford: Oxford University Press.

Suddaby, R., & Greenwood, R. (2005). Rhetorical strategies of legitimacy. *Administrative Science Quarterly, 50*(1), 35-67.

Tumasjan, A., Sprenger, T.O., Sandner, P.G., & Welpe, I.M. (2010). *Predicting elections with Twitter: What 140 characters reveal about political sentiment.* Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media.

Usherwood, S., & Wright, K. A. M. (2017). Sticks and stones: Comparing Twitter campaigning strategies in the EU referendum. *The British Journal of Politics and International Relations, 19*(2), 371-388.

Vaara, E. (2014). Struggles over legitimacy in the Eurozone crisis: Discursive legitimation strategies and their ideological underpinnings. *Discourse & Society, 25*(4): 500-518.

Verhaegen, S., & Hooghe, L. (2015). Does more knowledge on the European Union lead to more European identity? A comparative analysis among adolescents in 21 European member states. *Innovation: The European Journal of Social Science Research, 67*(2), 295-314.

Vliegenhart, R., Schuck, A.T., Boomgaarden, H.G., & De Vreese, C.H. (2008). News coverage and support for European integration, 1990-2006. *International Journal of Public Opinion Research, 20*(4), 415-439.

White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social structure from multiple networks. I. Blockmodels of roles and positions. *American journal of sociology, 81*(4), 730-780.

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of the conference on human language technology and empirical methods in natural language processing*, 347-354.

Appendix 1: Descriptive statistics of main variables

Q8_Good
Min.   :0.0000
1st Qu.:0.0000
Median :1.0000
Mean   :0.6174
3rd Qu.:1.0000
Max.   :1.0000

personalBenefit
Min.   :0.0000
1st Qu.:0.0000
Median :0.0000
Mean   :0.3283
3rd Qu.:1.0000
Max.   :1.0000

SimpleSentiment
Min.   :0.07224
1st Qu.:0.07224
Median :0.23494
Mean   :0.22857
3rd Qu.:0.46795
Max.   :0.46795

someMention
Min.   :0.000
1st Qu.:1.000
Median :2.000
Mean   :2.322
3rd Qu.:4.000
Max.   :4.000

proCap_comm_expense
Min.   :0.0000
1st Qu.:0.4939
Median :1.2929
Mean   :2.1875
3rd Qu.:3.6563
Max.   :5.3623

proCap_comm_expense2016
Min.   :0.000000
1st Qu.:0.000244
Median :1.369996
Mean   :1.341387
3rd Qu.:2.603094
Max.   :4.121650

SF_pc_m
Min.   : 15.37
1st Qu.: 67.60
Median :213.07
Mean   :175.28
3rd Qu.:248.88
Max.   :337.21

Q17
Min.   :1.000
1st Qu.:2.000
Median :2.000
Mean   :2.579
3rd Qu.:3.000
Max.   :4.000

Q9_3
Min.   : 0.00
1st Qu.: 5.00
Median : 7.00
Mean   : 6.58
3rd Qu.: 9.00
Max.   :10.00

Q9_2
Min.   : 0.000
1st Qu.: 6.000
Median : 8.000
Mean   : 7.596
3rd Qu.:10.000
Max.   :10.000

Q9_1
Min.   : 0.000
1st Qu.: 6.000
Median : 8.000
Mean   : 7.136
3rd Qu.: 9.000
Max.   :10.000

Q10_1
Min.   : 0.000
1st Qu.: 7.000
Median : 8.000
Mean   : 7.984
3rd Qu.:10.000
Max.   :10.000

Q10_2
Min.   : 0.000
1st Qu.: 3.000
Median : 6.000
Mean   : 5.761
3rd Qu.: 9.000
Max.   :10.000

Q10_3
Min.   : 0.00
1st Qu.: 4.00
Median : 7.00
Mean   : 6.06
3rd Qu.: 9.00
Max.   :10.00

Q10_4
Min.   : 0.000
1st Qu.: 2.000
Median : 5.000
Mean   : 5.344
3rd Qu.: 8.000
Max.   :10.000

Q10_5
Min.   : 0.000
1st Qu.: 5.000
Median : 7.000
Mean   : 6.293
3rd Qu.: 9.000
Max.   :10.000