

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

## **PERCEIVE**

**GA nr. 693529**

### **Deliverable 2.4**

# **Report on the probabilistic model of estimation of citizens' identification with the EU project and ranking of the case study regions**

Authors:

Valentina Aiello (valentina.aiello3@unibo.it)\*, Cristina Brasili (cristina.brasili@unibo.it)\*, Pinuccia Calia (pinuccia.calia@unibo.it)\*, Irene Monasterolo (irene.monasterolo@wu.ac.at)\*

\* Department of Statistical Sciences "Paolo Fortunati", Alma Mater Studiorum – University of Bologna

Leading partner in deliverable: Alma Mater Studiorum – Università di Bologna (UNIBO, Italy)

ESTIMATED PERSON/MONTHS PER PARTNER: UNIBO 7

OFFICIAL SUBMISSION DATE: 30, June 2018

ACTUAL SUBMISSION DATE: 29, June 2018

## Table of contents

<b>1. Introduction: aims, scope and <i>research question</i> of this deliverable.....</b>	<b>3</b>
<b>2. Operationalizing concepts and measures of European identity .....</b>	<b>4</b>
2.1 <i>Defining European identity: from theory into modelling .....</i>	4
2.1.1 From individual to European identity .....	5
2.1.2 Determinants of EU identification .....	8
<b>3. The <i>IdentEU</i> probabilistic model.....</b>	<b>10</b>
3.1 <i>Empirical model and strategy.....</i>	10
3.2 <i>Data and variables.....</i>	13
<b>4. Empirical analysis .....</b>	<b>16</b>
4.1 <i>Model specification and estimation .....</i>	16
4.2 <i>Description of clusters at the individual level .....</i>	17
4.3 <i>Identifying groups of regions .....</i>	22
4.4 <i>Mapping the level of identification with EU across European regions.....</i>	25
4.5 <i>Ranking of the case study regions.....</i>	28
4.6 <i>Effects of covariates.....</i>	32
<b>5. Concluding remarks .....</b>	<b>39</b>
<b>6. References.....</b>	<b>44</b>
<b>7. Appendix.....</b>	<b>47</b>

## **1. Introduction: aims, scope and *research question* of this deliverable**

In line with the Grant Agreement, this deliverable is provided on month 22 (June 2018) and aims to develop a novel probabilistic model to classify EU citizens according to their level of identification with the EU. In order to achieve this objective, D2.4 addresses the following research questions:

- To what extent do EU citizens identify with Europe and the EU project?
- Have European regions different patterns and level of identification?
- Are the results driven by specific socio-economic variables?

These research questions play a key role in the development of the PERCEIVE project's scientific results, embedding new challenges and opportunities in terms of methodological developments and policy implications at the EU level. On the one hand, D2.4 introduces an innovative conceptual and robust methodological solution to address the above-mentioned research questions, thus contributing to fill in a research gap. D2.4 helps us to contextualize and to better understand the current political context of the EU, which is characterized by growing Euro-skepticism and citizens' preference for populist parties, as well as by citizens' claims for democratization and transparency of the EU financial and economic decision making, at the light of EU citizens' identification with the EU institutions.

D2.4's research is well integrated and builds on the outcomes of previous PERCEIVE's work-packages activities, with the aim to exploit synergies and complementarities to deliver robust and policy-relevant results. In particular, D2.4 builds on the results of D1.1 (Regional case studies), on D1.2 (i.e. the PERCEIVE survey), D2.2. (determinants of citizens' identification with the EU and impact of the Cohesion Policy), D2.3 (Composite Index of European Identity), D5.1 (definition and the determinants of European identity), D5.3 and D5.4, as from Table A.1 (in Appendix). Then, D2.4 provides inputs for D2.5 and for D6.2-d.

The main outcome of D2.4 is the development of a novel probabilistic model - *IdentEU* - able to disentangle the patterns and drivers of EU citizens and regions' identification with the EU project. To our knowledge, this is the first model ever developed to address this crucial research question for the future of the EU. We opted for a Hierarchical Latent Class analysis model to produce a map of the level of identification with the EU project across European regions, through the classification of citizens and regions in clusters (6 and 4 respectively) identified by the classes of a latent variable, accounting for heterogeneity across regions both in terms of intensity of identification and of forms that it takes.

The results of the analysis developed in this report are supported by the development of the *IdentEU* probabilistic model and contribute to narrow the gap between academic research and policy making, in two ways. On the one hand, *IdentEU* contributes to fill a knowledge gap in terms of identification and understanding of the factors that influence citizens' attitudes toward the EU, both at individual and regional level. This information is crucial for EU policy makers in a historical period characterized by growing EU citizens' mistrust in the EU institutions, and growing Euro-skepticism, which eventually affected the recent electoral results in several EU member States (see Italy and France as most recent examples).

On the other hand, *IdentEU*'s results contribute to foster a better knowledge of the factors characterizing EU citizens' identification with the EU project. These elements provide policy-relevant and timely information to EU officers and policy-makers to improve the level of citizens' identification in the EU project. This is particularly timely and relevant in the context of the current discussion on the proposal for the programming period of EU Cohesion Policy 2021-2027.

Indeed, in the last decade, the economic and political elements gained importance in understanding defining citizens' EU identity and identification with the EU project. The negative effects of the last financial crisis on jobs and income accrued the effects of a mismanaged globalization process and eventually hit a large portion of EU population, increasing social and economic vulnerability for both low-income and middle-income households. The EU Cohesion Policy did not dispose of adequate financial endowments, nor was flexible enough to provide a buffer to the crisis and to move out of the crisis and identify opportunities for sustainable and inclusive growth in the EU. The worsening of the working and income conditions for a large share of EU households turned soon in social and political discontent. Therefore, although they are not explicitly considered in the model, in the conclusion we discuss the results obtained at the light of three phenomena that may have affected the citizens and regions' level of identification with Europe and the EU project, i.e. i) the effects of the 2008 financial crisis, ii) the process of globalization, and iii) the political turmoil generated by the Brexit, in the future research development ahead.

The report is organized as follows. Section 2 recalls the concepts and modelling approaches to measure EU citizens' identification with the EU at the light of geo-political and economic drivers, by building on previous work in sociology, political sciences and economics. The interdisciplinary background of the concept and determinants of EU identity, developed in previous PERCEIVE's work, represents our starting point to build the *IdentEU* model. Section 3 introduces the novel *IdentEU* probabilistic model to map EU citizens and regions' identification with the EU project, building on a Latent Class (LC) analysis and multilevel modelling to develop a Hierarchical Latent Class (HLC) model (Hagenaars and McCutcheon, 2002; Vermunt, 2003; Skrondal and Rabe-Hesketh, 2004; Pirani, 2013) and presents the data used for the analysis, coming from the PERCEIVE Survey (see D1.2), and secondary data including socio-economic variables at regional level. ). Section 4 presents the results of the model at the individual and regional level. Section 5 concludes providing policy-recommendations for the design of the EU Cohesion Policy framework 2021-2027.

## **2. Operationalizing concepts and measures of European identity**

### **2.1 Defining European identity: from theory into modelling**

In this section, we recall the main ideas and knowledge developed by previous PERCEIVE research results that are propaedeutic to build the *IdentEU* probabilistic model, which we describe in Section 3. In particular, we review the concepts and determinants of individual and collective identification with the EU and with the European project discussed in D5.1 (Barberio et al., 2017) and in D2.2 (Lopez-Bazo and Royuela, 2017) and used for the development of the Composite Index of European Identity in D2.3 (Royuela, 2018), identifying those functional to develop the probabilistic model *IdentEU* (see Section 3). In addition, we also identify and briefly discuss the socio-economic and political challenges that the EU has faced in the last decade and that should be considered in the discussion of the results of the probabilistic model *IdentEU*. The questions of why and to what extent EU citizens identify (or nor) with the EU institutions is gaining momentum due to the recent political developments in the EU, characterized by signals of "enlargement fatigue" and growing citizens' skepticism towards its institutions (see for instance the case of Brexit) and the Euro. The last electoral results in Italy and France, where populist parties and anti-system movements obtained remarkable results, are signal of an expanding gap between EU citizens and EU institutions. These trends have been analysed at the light of recent economic and

financial events both inside and outside the EU border, e.g. the role of an ill-managed globalization, the impact of the last financial crisis on wealth concentration and inequality, the European institutions' response to the Greek crisis, and from a governance point of view, highlighting the perceived lack of transparency in EU decision making in important areas of citizens' life, and the lack of a cohesive EU response to migration. The EU Cohesion Policy and its implementation at the regional, national and supranational level have been mostly absent from the debate. Nevertheless, the results of previous PERCEIVE analyses (see the focus groups at the regional level D1.1 and the answers to the PERCEIVE survey in D1.2) show that a large share of EU citizens is aware of the role of the EU Cohesion Policy and EU funds on their socio-economic conditions.

### 2.1.1 From individual to European identity

The concepts of identity and identification with the EU have been analysed in the literature in terms of individual and collective European identity.

The *individual* identity has been studied by Mendez and Batchler (2017) and Bergbauer (2018) at the light of the concept of social identity, defined as: “that part of the individual's self-concept which derives from their knowledge of their membership of a social group (or groups) together with the value and emotional significance of that membership” (Tajfel 1981 see also Tajfel and Turner, 1986). According to Bergbauer (2018), the individual identification with Europe can be explained as “citizens' self-categorisation as European together with their evaluations of their membership in the European collective and their affective attachment to Europe and other Europeans” (Bergbauer, 2018).

Three dimensions, i.e. a *cognitive*, *affective* and *evaluative* dimension, compose this subjective perception of identification (Bergbauer, 2018), where:

- the *cognitive* component refers to self-categorization as a member of a group (Awareness);
- the *affective* component refers to the emotional attachment and feeling of love and concern for the group, i.e. a “we-feeling” dimension (Attachment);
- the *evaluative* component refers to the assignment of value connotation (negative or positive) to the social group and his membership, by comparing people from the group with people out of the group. Mendez and Batchler (2017) link this dimension to the civic and cultural/ethnic distinctions of EU identity (Evaluation).

These dimensions have been also discussed in Royuela (2018) and used to inform the development of the Composite Index of European Identity.

The *collective identity* considers the presence of a feeling of solidarity within a group driven idea that a group of individuals accepts a central similarity, based on religion, ethnicity, language, social class, gender, and of course, nations (Royuela, 2018).

Bergbauer (2018) studies collective identity by building on concepts of social psychology, and on the sociological approach to collective identity. Social psychology considers the collective identity is “a situation in which individuals in a society identify with the collective and are aware that other members identify with this collective as well” (David and Bar-Tal, 2009, p. 361). The awareness of individuals identifying with the group allows the possibility of collective mobilization. In contrast, the affective ties (i.e. the we-feeling), and the level of political cohesion and solidarity between the members of a community (i.e. the sense of community) characterise the sociological approach (Easton, 1965). However, the aspects

related to the share and intensity of citizens' identification with Europe usually outpace those related to the level of awareness (Bergbauer, 2018).

Finally, we have the *European identity*, which has been studied by two streams of academic research, according to Hooghe and Verhaegen (2017), i.e.:

- The *society based-approach*, which assumes the identification of individuals with other European citizens and the trust towards other Europeans are respectively preconditions to the establishment of a European community, and to the legitimization of the process of European integration (Habermas, 2011; Risse, 2014);
- The *functionalist institutional approach*, which links the EU identification with the trust in the economic growth and prosperity and rights granted by European laws promoted by European institutions (Risse, 2010). This feeling has, in turn, both a cultural component referring to shared history, traditions and moral norms and values (Bruter, 2003), and a civic component, rooted on the rights and duties deriving from the EU Treaty (Reeskens and Hooghe, 2010).

Barberio et al. (2017), in D5.1 of the PERCEIVE project, build on Political Science, Social Sciences and International Relations to discuss the concept of EU identity, highlighting the institutionalist, the social identity, the integration, the citizenship and the spatial conceptualization of the discourse.

The *institutional approach* to the building of the EU identity emerged in the '60s of the last century and considered the introduction of a European identity as the final result of a process of European institutional building, which started with the European Economic Community and ended in the current form of the EU. It was expected that the benefits resulting from deeper cooperation among European member States would have increased trust in the European integration process, in line with the economic utilitarian theory (Gabel and Palmer, 1995).

Fligstein et al. (2012) link the *cooperation* approach and the *integration* approach, highlighting that the benefits of enhanced cooperation among European member States led to deeper integration. The different phases of EU integration contributed to the development of a nested structure of the concept of EU identity, starting from an initial economic integration phase to be followed by a political integration phase through spillovers, based on market mechanism (Haas, 1961). This process was expected to lead to a convergence of incomes, thanks to a period of long-lasting peace and cooperation in Europe, and through that to deliver co-operation, supranational rule-making and a convergence of beliefs, values and aspirations, generating a new European nationalism (Haas, 1968).

Nevertheless, the expected advantages of further integration have been recently questioned, leading to the development of a region-based and nation-based idea of EU integration. Indeed, the EU has experienced a growing diffusion of regionalist and separatist movements (the last being the Catalonia independence referendum in 2017), and the regional/national rhetoric of several Euro-skeptic parties. This trend was associated to the development of a language of regionalism and the territorialisation of space used by nationalists and regionalists in constructing spatial oppositions with the central authorities.

Then, in the late 1990s, the concept of social identities and the social construction idea emerged, contributing to shift the attention from rules and institutions to the social processes behind the construction of European integration and identification. According to Tajfel (1981), social identities are constructed in relation to, and in comparison with, other social identities, considering boundaries for inclusion and exclusion from the membership.

Barberio et al. (2017) also highlight the linkage between the European identity and social identity in so far "they refer to a dimension of the individual's self-concept shared with some



but not all other people” (Herrmann and Brewer, 2004; Citrin and Sides, 2004: 165). It emerges that social identities entail a locational and power dimension, where persons are associated with social expectations (Meyer and Hammerschmid, 2006).

More recently, the formation of European identity has been discussed in relation to that of *citizenship* and *national identities*, from the point of view of the Europeanisation of national identities. On the one hand, the concept of European identity has been discussed in terms of its alignment with the concept of *citizenship* (in so far it considers the presence of membership within a political community) and thus it is pivotal for integration. This means that, the more citizens identify with the EU, the more they support it.

On the other hand, EUROBAROMETER data showed that citizens who identify with their nation also have a basic sense of European identity (Citrin and Sides, 2004; Fligstein, 2008; Fligstein et al., 2012), with 43.3 percent of respondents viewing themselves as having primarily a national identity, but with European components.

At this regard, Marcussen et al. (1999) identify three conditions for the incorporation of the European dimension in the national identity, i.e. i) the *perception of legitimacy* of the new political order (which has to share elements with the previous national political order), ii) the *presence of critical junctures* when old ideas are being challenged or contested, and iii) a negative correlation between the *degree of consensus around national identities in a given society and the political elites’ space* to shape it. In addition, Risse (2010) defines Europeaness as a basis of national identity according to constellations of relations where i) identities can be nested, ii) identities can be cross-cutting (i.e. some, but not all, members of one identity group are also members of another one), and iii) a blending of identities where nestedness and cross-cutting identities are combined.

Finally, in a *functional* perspective, the discussion around the concept of EU identification questioned to what extent are EU institutions democratically legitimated (Habermas, 2011), analysing whether and how the EU institutions attempt to increase the democratic legitimacy of the multi-level political system (Holzhacker, 2007). This issue attracted more and more attention recently. Habermas (2012) discussed the issue of democratic legitimacy of EU institutions at the light of the management of the last financial crisis and of the Greek crisis, stating that in order for the European project to realize its democratic potential, it must evolve from an international into a cosmopolitan community, effectively extending the concept and implementation of democratic political institutions beyond the level of nation-states. Then, a discussion on the legitimacy of unelected institutions, i.e. the central banks, recently started (Tucker 2018). Indeed, central banks are unelected bodies but their power and influence on the economies increased considerably in the aftermath of the 2008 financial crisis, in particular with the introduction of unconventional monetary policies.

The answers to the PERCEIVE survey (D1.2) show that the pattern of citizens’ identification with the idea of Europe is very heterogeneous across and within member states, and thus far from being complete. Indeed, it was highlighted that the sense of loyalty to the political system emerges from the perception of the same political system being an effective problem-solving system (Braun and Tausendpfund, 2014; Harteveld et al., 2013).

These suggestions deriving from the survey will be analysed through the hierarchical latent class model (section 3) and the results of which will be analysed in detail (section 4).

### 2.1.2 Determinants of EU identification

The determinants of EU identity and individual identification have been discussed from a conceptual and an empirical point of view in Barberio et al. (2017), Lopez-Bazo and Royuela (2017), and Royuela (2018) and are operationalized in Royuela (2018). In Lopez-Bazo and Royuela (2017), they allowed to build a theoretical framework of analysis to understand the drivers and patterns of citizens' identification with the EU (using multivariate statistics and spatial analyses), while in Royuela (2018) they have been used to develop a Composite Index of European Identity aimed to assess the determinants of European citizens' identification with the European project (D2.3).

Barberio et al. (2017) discuss four determinants of EU identity from a conceptual point of view. They discuss the components of identity at the light of the dichotomy *civic vs. ethnic forms*, the *politicization* of European identity, and the *spatial* component, including the *rural-urban* divide.

First, the distinction between the *civic* form of identity, which focuses on citizenship as a legal status conferred by acceptance of common institutions, laws, rules, political and social systems, and the *ethnic/cultural* component of European identity, which is defined by the feeling of commonality with others based on a shared history, language, values, ideals and religion and it is mostly understood to be acquired per birth, has been widely addressed in the literature (see recent contributions by Bruter 2003; Reeskens and Hooghe, 2010; Risse, 2010). This means that while some citizens might identify more with the political and legal elements of the EU, others might identify more with the cultural elements of Europe. Royuela (2018) define the civic/ethnic divide referring to the Eurobarometer questions related to the drivers of EU citizens' feeling of community, in order to study whether a conflict exist between the two dimensions in relation to the EU identity. They use the answers to Eurobarometer's questions to create two indicators, one referring to the civic and the other to the ethnic dimensions of EU identity. The indicators are then used in a model of EU identity. The parameters of their aggregate model show that civic values are positively associated with European identity, while ethnic values show heterogeneous results, with a negative association of the index at the individual level (Royuela, 2018).

As regards the politicization of the European identity, the diffusion of Euro-skeptic and populist parties in the EU in the last 20 years (and in particular after the last EU Eastern enlargement in 2004, 2007 and 2013) contributed to develop the metaphor of the "Fortress Europe" (see Checkel and Katzenstein, 2009), whereas the idea of identity is based more on a cultural, religious and ethnic-self component hostile to foreigners, immigrants and refugees.

The spatial component plays a major role in the identification with the EU for at least two reasons (Paasi, 2001). On the one hand, in the last two decades, EU countries' political and social boundaries were relaxed, with the EU enlargement to Central and Eastern Europe, Baltic and Western Balkan countries, and with the freedom to move and work in the EU granted by the Schengen treaty. On the other hand, the responsibilities of regions and regional authorities kept increasing in several EU member states, as a consequence of the process of devolution of central governments' power, and the regionalization of the European Cohesion Policy. In addition, the access to the common EU market led to a globalization of production and consumption patterns, and to the loss of regional and local specificities (even in areas characterised by traditional productions, such as the industrial districts). In particular, the analysis of the drivers of the relation between urban and rural areas have shaped convergence studies in the last two decades.

Lopez-Bazo and Royuela (2017) review the discourse on the determinants of EU identification at the individual and the system level, as discussed by Bergbauer (2018). The



*individual* level is characterized by political awareness, attitudes towards the European and national bodies, and personal transnational experiences. In contrast, the *system* level is influenced by party messages related to European and national community, the economic position and degree of international integration, and the ethnocultural identification (see Royuela, 2018). Then, based on the literature review, they provide a taxonomy of seven main determinants of identification with the EU, i.e.:

- *Political awareness*, which considers the political interest and knowledge in general and EU matters, affecting the identification with Europe based on the information on the benefits (risks and downsides) of European integration.
- *Attitude towards the EU and national bodies*, where the strength of national identification affects the level of identification with the EU, in a positive way if we accept the possibility of complementarity of identities, or in a negative way, if concerns on integrity and sovereignty drive to conflicting identities.
- *Personal traits and experience*, where personal contacts with other EU citizens, including cross border trips or living together with intra-European immigrants, are expected to increase individual identification with Europe.
- *Party message related to the EU and national community*, where the prevalence of pro-EU parties in a country lead to higher citizens' exposure to messages signalling the benefits of the European integration (and viceversa, in the case of prevalence of Euro-skeptic countries).
- *National economic position and degree of international integration*. This pertains citizens' perceived benefits of belonging to the EU, and in particular, as analysed in other PERCEIVE research (e.g. D1.1, D1.2, D2.2, D5.1), in relation to the communication of the results of the Cohesion Policy at the national, regional and local level.
- *Ethnocultural identification*, which can assume both a positive connotation, when citizens perceive that they gain experiences by interacting with other Europeans, and a negative connotation, when citizens perceive only the costs associated with further EU integration e.g. on changes in the labour market due to migration.
- *Trust in political institutions*, at different levels, i.e. regional, national and European level. There is growing evidence that trust in institutions at the national level is positively correlated with trust in EU institutions (Hooghe and Verhaegen, 2017).

From an empirical point of view, the identification with the idea of Europe has been represented by a wide list of determinants and translated into concrete indicators of the dimensions proposed by the theoretical approaches. In particular, Lopez-Bazo and Royuela (2017) identify several individual control variables used in the literature, covering "age, gender, ethnicity, religion, education, occupation, the perceived situation of the economy and or society (e.g. perceived financial situation of the household) and even psychological traits, such as life satisfaction, partly (or even mostly) driven by external circumstances" (Lopez-Bazo and Royuela, 2017). Main influencing variables are grouped into variables referring to:

- *Political capital*, and including "cognitive mobilization, satisfaction with domestic democracy, perceived benefits from EU membership, and trust in institutions (although this variable is also considered as a proxy to identification with such institutions)" (Lopez-Bazo and Royuela, 2017)
- *Ideology*, in particular referring to the political orientation and to the democratic satisfaction at the EU and country level.

- *Socio-economic context*, including control variables (e.g. GDP growth, inflation and unemployment, Verhaegen et al., 2014)); the rural - urban divide (Luhman, 2017); old versus new EU member states or years of EU membership.
- *Country level variables*, including the corruption index, the scope of the welfare state and the economic benefit of belonging to the EU (e.g. the net contribution to EU budget, received structural funds, spread on sovereign bonds and intra EU export). In particular, the extent to which governments strategically allocate regional transfers to influence the public opinion has been analysed, also in relation to the effect of the EU structural funds on citizens' attitude towards the EU (Osterloh, 2011).

A specific discourse is then developed in Lopez-Bazo and Royuela (2017) and Royuela (2018) on the role of the EU Cohesion Policy on EU identification. Lopez-Bazo and Royuela (2017) discuss Osterloh's results (2011), which highlight that EU Regional Policy affects citizens' awareness and thus their support to the EU, but the level of awareness and support is conditioned to socio-economic characteristics, such as education. Levels of individual awareness are heterogeneous within every country, and relevant differences exist between regions in terms of education inequality (Rodriguez-Pose and Tselios, 2009, 2010) and urban and rural divide (Rodriguez-Pose and Tselios, 2011).

The review of the discussion on the determinants of EU identification is propaedeutic for the development of the *IdentEU* probabilistic model described in Section 5. Indeed, the model integrates the three dimensions of "Awareness" (in terms of the EU Cohesion Policy and its local benefits, as well as their level of identification with Europe, country, region, and European values), "Evaluation" (in terms of perceived effectiveness of political institutions and benefits of EU membership) and "Attachment" (i.e. participation in the election and redistributive role of the EU). In addition, *IdentEU* considers: i) demographic and socio-economic characteristics, introduced in the model as control variables and predictors (i.e. age, education, occupation status and the economic status, see table A.3 in appendix for a comprehensive description), ii) the spatial dimension, considered in the contextual covariates (e.g., the level of GDP per inhabitant in Euro at 2014 as percentage of EU average, the unemployment rate for people of 20-64 years old at 2014, the absorption rate of the Structural Fund financial allocation at 2013).

### **3. The *IdentEU* probabilistic model**

#### **3.1 Empirical model and strategy**

The analysis of the nature of what influences identity, and its implications in terms of attitudes towards integration and questions of behavior, have been mostly studied by political scientists with qualitative models (Meinhof and Galasinski, 2005; Grundy and Jamieson, 2007; Bruter, 2004). Herrmann et al. (2004) and Bruter (2005, 2009) discussed how to move from qualitative to operative quantitative models, eventually identifying several challenges related to the measurement and characterization of identity. In particular, Bruter (2008) pointed out how in order to answer some major paradoxes, a more critical and rigorous measurement of European identity is needed. Bruter (2008) asserts that "most of the models that conclude to widespread Euroscepticism and minimal European identity rely on measures which truly capture neither". Thus, imprecise or inaccurate measures of European identity coming from Eurobarometer's Moreno questions can only lead to biased models. This explains the poor literature on quantitative modeling of EU identity, at the aggregate and

individual level. The first empirical model of EU identity, which relies on Herrmann et al. (2004) analysis of coexistence of multiple identities, was provided by Bruter (2009) who uses an experimental panel study design to test dynamic hypotheses, and finds a positive correlation between citizens' European and national identities of the order of 0.19.

With this deliverable, we contribute to enrich the infant literature on quantitative modelling of EU identity by developing the *IdentEU* probabilistic model. We opted for Latent Class Analysis (LCA) and multilevel modelling to develop a Hierarchical Latent Class (LC) model (Hagenaars and McCutcheon, 2002; Vermunt, 2003; Skrondal and Rabe-Hesketh, 2004).

Latent Class Analysis (LCA) is a powerful tool for the analysis of the association among categorically observed measures (*indicators* or *manifest variables*) with the aim of identifying a latent categorical variable accounting for the interrelationships and the associations between the observed indicators. The latent classes are the levels of this categorical latent variable, to which the respondents are allocated depending on the responses to the observed indicators. Hence, LCA can serve to investigate the unobserved heterogeneity among the respondents and to develop a probabilistic classification of respondents on the basis of their responses.

Through LCA, we deal with identification of citizens with EU as latent construct. In our empirical application, the latent construct is derived from a set of interrelated observed measures (described in the next section) whose associations are due to an unobserved factor treated as categorical: the latent classes represent the latent levels of identification.

Despite being possible to consider identification as a latent continuous variable, assuming a discrete latent variable allows us to define different patterns of identification according to the different profiles of respondents, and to analyse their features.

We start from a set of  $K$  categorical observed indicators<sup>1</sup>. The model identifies  $T$  classes of a latent variable that describes an unobservable construct (i.e. the identification with the EU) and provides a classification of individuals based on the response patterns to the  $K$  indicators. In doing so, it exploits the nested structure of the data, insofar the individuals (first-level units) are nested into regions (second-levels units). Our hierarchical model accounts for unobserved (latent) regional effects and considers them as a discrete latent variable with  $L$  classes according with the classification of identification at individual level.

From a methodological point of view, we indicate the  $K$  responses referring to the concept of identification with the EU with  $Y_{ijk}$ , each describing the response to item  $k$  of person  $i$  coming from region  $j$ .

In particular:  $\mathbf{Y}_{ij}$  is the vector of responses of the same individual  $i$ ,  $\mathbf{Y}_j$  represents the vector of responses of all individuals in region  $j$ ,  $\mathbf{Z}_{ij}$  is the vector of individual covariates, and  $\mathbf{Z}_j^g$  is the vector of covariates at regional level. Then, we define as  $X_{ij}$  the first-level latent variable (unobservable) that represents identification with EU at individual level.

Given their response patterns to the manifest indicators, individuals are classified in one of the  $T$  latent classes of  $X_{ij}$ ,  $t=1, \dots, T$ , where each latent class is identified by the pattern of the  $K$  individual responses classified with the highest probability in that class (Standard LC Model).

The random effects at regional level are specified as a discrete latent variable  $W_j$ , conditionally on which the individual responses are assumed to be mutually independent.  $W_j$  identifies latent types of regions for which the parameters in the model differ. Thus, the second level latent variable allows to cluster the second level units (i.e. regions) into a small number of latent classes,  $m=1, \dots, M$ .

---

<sup>1</sup> A similar model is discussed in Pirani (2013) and used to investigate the patterns of social exclusion in the European regions. In describing the model, we follow Pirani and the notation used there.

LC models are based on the assumption of “local independence”, which implies that the response variables are mutually independent given the latent variable. This means that within the latent classes, only random relationships among variables remain, and LCA assigns the units into latent classes so that the indicators are uncorrelated within each class.

Formally, the local independence assumption implies that the observed indicators  $Y_{ijk}$  are mutually independent given the latent variable, i.e. they are connected only indirectly through the common latent variable  $X_{ij}$ . The latent variable is assumed to explain all the associations among the manifest variables. Standard LCA determines the smallest number of latent classes sufficient to account for the associations among the manifest variables. Given the local independence assumption, the probability of observed data can be written as:

$$P(\mathbf{Y}_j | \mathbf{Z}_j) = \sum_{m=1}^M \left[ P(W_j = m | \mathbf{Z}_j^g) \left[ \prod_{i=1}^{n_j} \sum_{t=1}^T P(X_{ij} = t | W_j, \mathbf{Z}_{ij}) \prod_{k=1}^K P(Y_{ijk} = s_k | X_{ij}, W_j) \right] \right]$$

where three components can be identified, each specified using a multinomial logit specification, as:

- 1) the latent class probability at regional level;
- 2) the latent class probability at individual level;
- 3) the conditional probability of individual response pattern.

The latent class probability at regional level is the probability of region  $j$  belonging to a particular class of the latent variable  $W_j$ , given the regional covariates:

$$P(W_j = m | \mathbf{Z}_j^g) = \frac{\exp(\alpha_{0m} + \sum_l \alpha_{lm} Z_{lj}^g)}{\sum_{m'=1}^M \exp(\alpha_{0m'} + \sum_l \alpha_{lm'} Z_{lj}^g)}$$

where  $l$  is an index for the group-level covariates. It is assumed that the probability of latent class membership at regional level depends on the category effect of the latent variable at regional level  $\alpha_{0m}$  and on the effects of the contextual (group level) covariates on  $W_j$ . These probabilities provide information about the distribution of the population among the regional classes.

In contrast, the latent class probability at individual level is the probability that the respondent  $i$  of the  $j$ -th region belongs to a particular class of the individual level latent variable  $X_{ij}$ , given regional latent class membership and the individual covariates:

$$P(X_{ij} = t | W_j = m, \mathbf{Z}_{ij}) = \frac{\exp(\gamma_{0tm} + \sum_l \gamma_{lt} Z_{lij})}{\sum_{t'=1}^M \exp(\gamma_{0t'm} + \sum_l \gamma_{lt'} Z_{lij})}$$

where  $l$  is an index for the individual-level covariates. It is assumed that the probability of belonging to a certain individual level latent class depends on the group-level latent variable ( $\gamma_{0tm}$  captures the differences between the classes of  $W_j$  in the category effect of  $X_{ij}$ ) and on the individual level covariates. These provide information about the distribution of the population among the individual classes.

Finally, the conditional probability of individual response pattern is the joint probability of the  $i$ -th respondent following the pattern  $s_i$  given individual and regional latent class membership:

$$\prod_{k=1}^K P(Y_{ijk} = s_k | X_{ij} = t, W_j = m) = \prod_{k=1}^K \frac{\exp(\beta_{0s_k} + \beta_{1s_k t} + \beta_{2s_k m})}{\sum_{s'=1}^{S_k} \exp(\beta_{0s'} + \beta_{1s' t} + \beta_{2s' m})}$$

where  $\beta_{0s_k}$  represents the specific effect of the response category and  $\beta_{1s_k t}$  is the main effect of the individual latent variable  $X_{ij}$ . According to the different indicators, conditional probabilities can be modelled in different ways assuming also direct effects of the group-level latent variable  $\beta_{2s_k m}$ . In this case both the individual level class proportion and the class-specific probabilities depend on  $W_j$ .

These probabilities provide information that is useful to describe the latent classes, and to obtain a profile for the latent classes according to which responses are prevalent in each of them.

### 3.2 Data and variables

In this section, we describe the data and variables used to develop the *IdentEU* probabilistic model aimed to classify individuals and regions according to their identification with the EU project on the basis of the results of previous deliverables (see Section 1) and from the literature review (see section 2).

In particular, for building the model we rely on two sources of data:

- The PERCEIVE Survey (D1.2);
- A dataset collecting socio-economic variables at regional level developed within the PERCEIVE project (D2.1).

The PERCEIVE Survey was designed by the University of Gothenburg as a specific task of the PERCEIVE project. The survey collects information from a sample of 17.147 individuals of 18 years of age or older from 15 EU member States and was conducted during the summer of 2017. The survey design and preliminary results are described in Bauhr and Charron (2018).

The survey includes 35 questions regarding different aspects, focusing on: i) respondents' awareness of EU Regional Policy, ii) their identification with Europe, country, region, and European values, iii) their Political attitudes and values. In addition, the survey includes questions aimed to measure the respondents' support for the Cohesion Policy, as well as their demographic and socio-economic characteristics. Table 1 shows the composition of the sample by country.



**Table 1.** Sample composition: number of respondents, their percentage distribution, and number of regions by country

Country	Number of respondents	Percentage distribution	Number of regions
<i>France</i>	1,500	8.75	22
<i>Bulgaria</i>	503	2.93	6
<i>Slovakia</i>	1,014	5.91	4
<i>Hungary</i>	1,000	5.83	7
<i>Romania</i>	1,015	5.92	8
<i>Italy</i>	2,000	11.66	17
<i>Netherlands</i>	500	2.92	14
<i>Sweden</i>	580	3.38	3
<i>UK</i>	1,500	8.75	12
<i>Latvija</i>	500	2.92	1
<i>Poland</i>	2,000	11.66	16
<i>Spain</i>	2,014	11.75	17
<i>Germany</i>	1,500	8.75	16
<i>Estonia</i>	521	3.04	1
<i>Austria</i>	1,000	5.83	9
<b>Total</b>	17,147	100	153

The regional aggregation in the survey is at level NUTS2 for the majority of the countries in the sample, except for Germany, UK and Sweden (level of aggregation at NUTS1), and for Latvija and Estonia (the whole country).

The second source is a dataset that brings together information from several official secondary sources at regional level (NUTS1 and NUTS2) in a panel format (Charron, 2017). We use the PERCEIVE survey's questions to select the manifest variables as proxies to the latent construct of identification with EU, i.e. the  $K$  manifest variables labelled  $Y_{ijk}$  in the methodological section. We start from the set of variables used in Royuela (2018) in building the synthetic indicator for identification with EU. These variables allow us to represent the three components of the concept of individual identification described in section 2, i.e. "awareness", "evaluation", "attachment".

As regards the component "awareness", we use the responses to the questions:

Q1. Have you ever heard about the following EU policies? (EU Cohesion Policy; EU Regional Policy; EU Structural Funds; any EU funded project in your region or area)

We recoded the responses to build up a variable with three categories: 0 -None of them; 1 – Only local project; 2 – At least one.

The category "Only local project" includes people who responded they know only EU funded project in their region or area, while the category "At least one" includes people who know at least one among Cohesion Policy, EU Regional Policy, and EU Structural Funds.

Q3. Have you ever benefited in your daily life from any project funded by the EU?

This is a binary variable with the following categories: 1 – Yes; 2 - No

Moreover, we include the responses to the question regarding the strength of identification with Europe, Country, and Region:

Q9. On a 0-10 scale, with ‘0’ being “I don’t identify at all” and ‘10’ being “I identify very strongly”, how strongly you identify yourself with the following: Q9\_1: Your region; Q9\_2: Your country; Q9\_3: Europe?

The original scale would lead to consider too many categories, implying a very large number of parameters in the model and consequently possible estimation problems, hence we decided to group the values in a smaller number of categories. To this end, we group the responses in the original scale into three categories of identification level: 1 – Not strongly (including the values 0, 1, 2, 3); 2 – somewhat strongly (including values 4, 5, 6); 3 – Strongly (including values 7, 8, 9, 10). We opted for this classification on the basis of the distribution of the original responses, which shows a concentration in the central and extreme categories, while respondents are more uniformly distributed in the intermediated ones; moreover it solve the problem of where to assign the central value in the case of an even number categories.

We consider one variable describing the level of identification with EU and build other two variables comparing the level of identification in EU to the level of identification with the country and the region respectively:

- Identification with EU: 1 – Not strongly; 2 – Somewhat strongly; 3 - Strongly
- Identification with EU vs Country: 1 – Less; 2 – Equal; 3 – More
- Identification with EU vs Region: 1 – Less; 2 – Equal; 3 – More

As regards the component “Evaluation”, we use the responses to the questions:

Q5. How effective do you think the following institutions will be at dealing with the biggest problem in your region? (Q5\_1: The European Union; Q5\_2: National governing institutions; Q5\_3: Regional/local governing institutions)

Then, moving from the responses to Q5 (but inverting the ordering of the scale), we build the following three variables:

- Effectiveness of EU: 1: Not so effective; 2 – Somewhat effective; 3 – Very effective
- Effectiveness of EU vs National institution: 1 – Less; 2 – Equal; 3 – More
- Effectiveness of EU vs Regional/local institution: 1 – Less; 2 – Equal; 3 – More

Q8. In general, do you think that (YOUR COUNTRY’S) EU membership is a good thing, a bad thing, neither good nor bad?

This question was not posed to UK respondents due to the recent experience of the Brexit referendum at the time the Survey was carried out. We construct a variable grouping the responses into two categories: 1 – “Good thing”; 0 – “Bad / Neither good nor bad”. We imputed UK respondents the value 1 - “Good thing” if they answer they would have voted “remain” if the Brexit referendum were held again at the time of the survey, and 0 otherwise.

Q16. On a 0-10 scale, with ‘0’ being that ‘there is no corruption’ and ‘10’ being that corruption is widespread, how would you rate: Q16.1: the European Union; Q16.2: the national government; Q16.3: the region/local government?

We build three variables, grouping the responses in the original scale into three categories of corruption level: 1 – Low (including the values 0, 1, 2, 3); 2 – Medium (including values 4, 5, 6); 3 – High (including values 7, 8, 9, 10).

The first variable we consider regards the perceived level of corruption in EU, the other two are obtained comparing the perceived level of corruption in EU to the perceived level of corruption in national government and regional government respectively:

- Corruption in EU vs National government: 1 – Less; 2 – Equal; 3 – More
- Corruption in EU vs Regional/local government: 1 – Less; 2 – Equal; 3 – More

As regards the component “Attachment”, we use the responses to the question:

Q7. Have you voted in either of the last two EU parliamentary elections?

We use a variable with the following categories: 0 – Neither; 1 – Once; 2 – Both times

Moreover, we decided to include another indicator to describe the component “attachment” using the responses to the question:

Q20: In your opinion, the EU should continue this policy, where wealthier countries contribute more, and poorer EU regions receive more funding?

From the original five classes (Strongly agree, Agree, Disagree, Strongly disagree, D/K) we obtain 3 classes: 1 – Agree; 2 – Disagree; 3 – D/K. This indicator measures the citizen support to the values put forward by EU, hence it may be used to proxy their attachment to Europe and other Europeans.

A descriptive analysis regarding each of these variables (and other information from the survey) are reported in Bauhr and Charron (2018), as well as in previous deliverables (Lopez-Bazo and Royuela, 2017; Royuela, 2018); in particular, in Lopez-Bazo and Royuela (2017) there is also an analysis of the distribution of each indicator at the aggregate regional level.

As for the variables to be included in the model as controls and predictors, we consider the individual characteristics as age, education, occupation status and the economic status (see table A.2 in appendix for definition and categories). Contextual covariates are introduced at the regional level. These regard the level of GDP per inhabitant in Euro at 2014 as percentage of EU average, the unemployment rate for people of 20-64 years old at 2014, the per capita expenditure of Structural Fund in the whole programming period 2007-13, the European Index of Institutional Quality (EQI) at 2013 (normalised at 100), and the absorption rate of the Structural Fund financial allocation at 2013.

While the EQI assesses the institutional quality based on the experiences and perception of citizens (Charron et al., 2014), the absorption rate could be considered as an objective measure of the regions’ effectiveness and efficiency in programming and implementing Cohesion Policy. Several studies show (Bojmans, 2014; Tosun, 2013) a strong correlation between absorption rate and institutional quality at country level.

## 4. Empirical analysis

### 4.1 Model specification and estimation

The model estimation requires us to fix the number of classes at the group level and at the individual level, i.e.  $M$  and  $T$  respectively. The final model we present involves 6 latent classes at individual level,  $T=6$ , and 4 classes at regional level,  $M=4$ <sup>2</sup>.

The choice of the number of individuals’ clusters and groups of regions is determined by the comparison of alternative model specifications, and by selecting among models by the minimum BIC (Bayesian Information Criterion) rule (Table A.3 in appendix shows the BIC values for different models). For the choice of the final model, we also consider other

---

<sup>2</sup> For model estimation we use the software LATENT GOLD 5.1 (Vermunt and Magdison, 2016)

measures that are useful for the interpretation of the model, such as bivariate residuals, classification errors and the profile of the different classes associated to the model.

Bivariate residuals (BVR) are useful to check for the presence of residual association between pairs of variables after the model estimation. During the phase of model selection, the analysis of residuals helped in selecting the most appropriate indicators, as well as the number of latent classes to retain.

Among the indicators presented in Section 3, some of them have been discarded from the final analysis because high residual associations after model estimation indicate redundant information. In particular, the indicator obtained from question Q3 (Have you ever benefited in your daily life from any project funded by the EU?) is redundant with respect to the information brought by the indicator “Awareness of Cohesion Policy”, as well as all the indicators measuring the level of identification with Europe, and EU effectiveness and corruption in comparison with regional/local level institutions that are high correlated with the same variables measured at the national level.

Inspection of bivariate residual during the model selection process, moreover, can provide information for the need of relaxing the local independence assumption and considering direct effects of level-2 latent variable and covariates on indicators, or between indicators themselves. Hence, in the specification of the final version of the model, we also make the following assumptions:

- The existence of a direct effect of the latent variable at regional level  $W_j$  on the indicator “Evaluation of EU membership” (from question Q8), which measures how individual evaluate the EU membership of their country. High residual association (not explained by the model) of this indicator with many contextual covariates at regional level has lead us to suppose that the regional contextual environment can probably have a direct influence on whether people think the EU membership is a good o a bad thing.
- The existence of a direct effect of the amount of Structural Funds (SF) expenses at regional level (per capita SF expenses in the region in the programming period 2007-13) on the indicator “Awareness of Cohesion Policy”, which describes to what extent individuals know Cohesion Policy; here the rationale is that people more likelihood know Cohesion Policy if they live in regions where the amount of EU financial funding is higher.

Some associations (high residuals) between certain indicators, and between covariates and indicators, still remain after controlling for the latent variables. However, relaxing conditional independence assumption for all these cases would increase too much the number of parameters, then implying computational problems. Moreover, increasing the number of latent classes does not improve significantly neither the model fitting nor its interpretation. Thus, the choice of the final model specification was based on the interpretability of the latent classes too.

We comment the results in 4.2 by looking at the estimated probabilities.

## 4.2 Description of clusters at the individual level

We consider the classification of individuals based on the *posterior class membership probability*. This classification gives information on how well one can predict to which latent class individuals belong given their observed indicators and covariates patterns. For each

subject  $i$  with certain covariates and response pattern  $\mathbf{s}$ , the probability of belonging to the latent class  $t$  of  $X_i$  can be obtained by means of the Bayesian rule:

$$\hat{P}(X_i = t|Y_i = \mathbf{s}, \mathbf{Z}_i) = \frac{\hat{P}(X_i = t|Y_i = \mathbf{s}, \mathbf{Z}_i)\hat{P}(Y_i = \mathbf{s}|X_i = t, \mathbf{Z}_i)}{\hat{P}(Y_i = \mathbf{s}|\mathbf{Z}_i)}$$

where the numerator and the denominator are the Maximum Likelihood estimates.

The most common classification rule consists to assign each individual to the latent class with the highest  $\hat{P}(X_i = t|Y_i = \mathbf{s}, \mathbf{Z}_i)$  (modal assignment). The classification table (Table 2) cross-tabulates posterior and modal class membership probabilities (Vermunt and Magidson, 2016). Each entry in the table represents the sum of the class  $t$  posterior membership probabilities for the cases allocated to each modal class. The off-diagonals cases are the misclassified ones, underlying which latent classes are well separated.

**Table 2.** Classification table of latent variable  $X_{ij}$  based on posterior class membership probabilities  $\hat{P}(X_i = t|Y_i = \mathbf{s}, \mathbf{Z}_i)$

Probabilistic	Modal assignment						Total
	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	
Cluster1	4246.4	117.1	0.3	186.9	58.7	13.5	4622.7
Cluster2	82.9	3012.1	0.0	1.7	181.9	86.8	3365.4
Cluster3	0.2	0.0	2960.8	0.2	46.9	17.2	3025.3
Cluster4	331.7	3.7	0.0	2113.4	4.4	66.6	2519.9
Cluster5	58.0	127.0	35.4	0.3	1473.5	83.9	1778.1
Cluster6	18.8	65.2	41.5	44.4	79.7	1428.0	1677.5
Total	4738.0	3325.0	3038.0	2347.0	1845.0	1696.0	16989.0

The classification error, ranging from 0 to 1, is the proportion of misclassified cases and it helps to evaluate the distinctiveness of different classes. This proportion indicates how well the model can predict latent class membership given the value of indicators and covariates. Concerning our final model specification, the classification error of individuals is 0.10. Table 2 shows that Cluster 1 is well distinct from Clusters 3, 5 and 6; Cluster 2 is well distinct from Clusters 3, 4 and 6; and Cluster 3 is well distinct from Clusters 4, 5 and 6. In contrast, misclassification happens between Clusters 1 and 4, which share similar aspects, and to a lesser extent between Cluster 2 and Clusters 1 and 5.

The profile table (Table 3) shows the characteristics of each cluster, similarities and differences. In the first row, the *estimated marginal latent probabilities*  $\hat{P}(X = t)$ , obtained by aggregating the model probabilities  $\hat{P}(X_i = t|\mathbf{Z}_i)$  over covariates values when the model contains covariates (Vermunt and Magidson, 2016), are reported.

The other values represent the *class-specific marginal probabilities* associated with each indicator  $P(Y_{ijk} = s_k|X = t)$  describing how the latent classes are related to the indicator variables. These probabilities sum to 1 within each class. By analysing the profile table, we can characterize and name each class of the latent variable in terms of response probability to each level of the indicators<sup>3</sup>.

<sup>3</sup> In LCA models with direct effects of covariates on indicators and direct association between indicators, probabilities should be obtained by aggregating, for each indicator, over the other variables involved in the effects specification. Also, in multilevel models with direct effect of group level latent variable on one or more



In the final model, we identify six clusters of respondent types. The clusters are ranked by their size, from the largest to the smallest: Cluster 1 is composed by 27% of individuals, Cluster 2 by about 20% and the smallest is Cluster 6 comprises 10% of individuals.

Looking at marginal probabilities for each indicator, we clearly identify three clusters of individuals that, more than others, strongly identify with Europe, and three clusters of individuals that, more than others, do not considerably identify with Europe. The probability to identify strongly with Europe is very high, about 0.9 or more, for individuals in Clusters 4, 1 and 3. On the contrary, individuals in Cluster 2 have the highest probability to respond they do not identify at all or not too much (0.46). Individuals in Cluster 5 and 6 have very small probability to answer that they strongly identify with Europe but have a higher probability to identify somewhat stronger compared to Cluster 2.

Regarding the second indicator in the table (*Europe vs Country identification*), which measures the strength of individual identification with Europe compared to his own country, we observe that individuals in Clusters 1, 3 and 4 identify with Europe as strongly as they identify with their own country, while individuals in Clusters 2, 5 or 6 have the highest chance to identify less with Europe than with their country.

However, we find differences among clusters inside each of these two groups when considering how they evaluate the effectiveness in solving problems and the level of corruption of EU institutions.

---

indicators, the marginal probabilities are obtained summing over the classes of this latent variable (see Vermunt and Magidson 2016 for further details).

**Table 3.** Profile table of the latent variable at individual level  $X_{ij}$ : size class  $\hat{P}(X_i = t|Z_i)$  and class specific marginal probabilities  $P(Y_{ijk} = s_k|X = t)$

	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>	<b>Cluster 4</b>	<b>Cluster 5</b>	<b>Cluster 6</b>
	<i>Disappointend pro- Europe</i>	<i>EU Deniers</i>	<i>Confident Europeans</i>	<i>Wary pro- Europe</i>	<i>Disaffected Europeans</i>	<i>Wary Cons- Europe</i>
<b>Cluster Size</b>	0.2727	0.1972	0.1752	0.1495	0.1056	0.0999
<b>Indicators</b>						
<b><i>How strongly identify with Europe</i></b>						
<i>Not much strongly</i>	0	0.4612	0.0362	0	0.2704	0.2857
<i>Somewhat</i>	0.0995	0.5387	0.0681	0.0726	0.7293	0.7142
<i>Strongly</i>	0.9005	0	0.8957	0.9273	0.0003	0.0001
<b><i>Europe vs Country identification</i></b>						
<i>Less</i>	0	0.6849	0	0	0.6716	0.7187
<i>Equal</i>	0.8595	0.2942	0.8271	0.8557	0.3193	0.2573
<i>More</i>	0.1405	0.0209	0.1729	0.1443	0.0091	0.0241
<b><i>Effectiveness in solving problems</i></b>						
<i>Not so Effective</i>	0.7413	0.9007	0.3498	0.0001	0.7966	0.0001
<i>Somewhat effective</i>	0.2587	0.0957	0.4656	0.6501	0.2033	0.7156
<i>Very effective</i>	0	0.0036	0.1845	0.3498	0	0.2843
<b><i>EU vs National effectiveness</i></b>						
<i>Less</i>	0.3386	0.3069	0.1694	0.0071	0.3654	0.0085
<i>Equal</i>	0.6614	0.6931	0.5138	0.4274	0.6346	0.4143
<i>More</i>	0	0	0.3168	0.5655	0	0.5772
<b><i>Corruption in EU</i></b>						
<i>Low</i>	0.082	0.0006	0.3341	0.0846	0.2302	0.1165
<i>Medium</i>	0.3462	0.1368	0.6659	0.2663	0.7697	0.3861
<i>High</i>	0.5718	0.8625	0	0.6491	0.0001	0.4974
<b><i>EU vs National Corruption</i></b>						
<i>Less</i>	0.0001	0	0.9998	0	0.5253	0.255
<i>Equal</i>	0.7792	0.729	0.0002	0.8458	0.4226	0.6135
<i>More</i>	0.2207	0.2709	0	0.1542	0.0522	0.1315
<b><i>Vote</i></b>						
<i>Neither</i>	0.2869	0.3916	0.2533	0.2642	0.3904	0.3718
<i>Once</i>	0.1481	0.1433	0.168	0.1755	0.1924	0.1797
<i>Both times</i>	0.5468	0.4456	0.5561	0.5423	0.3945	0.4181
<i>(d/k-refused)</i>	0.0182	0.0195	0.0227	0.018	0.0226	0.0304
<b><i>Support to Cohesion policy</i></b>						
<i>Agree</i>	0.8291	0.6238	0.9022	0.8845	0.7556	0.8007
<i>Disagree</i>	0.1576	0.3642	0.087	0.1061	0.2271	0.1832
<i>d/k</i>	0.0133	0.012	0.0108	0.0094	0.0173	0.0162
<b><i>EU membership</i></b>						
<i>Bad thing</i>	0.2751	0.7464	0.1743	0.1727	0.4786	0.4718
<i>Good Thing</i>	0.7249	0.2536	0.8257	0.8273	0.5214	0.5282
<b><i>Awareness of Cohesion policy</i></b>						
<i>None</i>	0.1819	0.2664	0.1019	0.1141	0.2578	0.2082
<i>only local project</i>	0.0863	0.1053	0.1291	0.1023	0.1078	0.1236
<i>Cohesion/regional policies</i>	0.7318	0.6283	0.769	0.7836	0.6344	0.6683

Actually, individuals in Cluster 1 evaluate the effectiveness of EU in solving the problems more negatively than others, judging it even less effective or at same level than the national governing institutions, and they perceive a high level of corruption in EU institutions, equal or more widespread than in national institutions. Regarding the issue of redistribution, according to which richer countries should contribute more to the EU Cohesion Policy and poorest regions should receive more, individuals in Cluster 1 “Disappointed pro-Europe” agree with high probability (even if this is not the highest across clusters). Moreover, with high probability, they evaluate the EU membership as a good thing (though the probability is the lowest among Clusters 1, 3 and 4). The probability that people know the existence of Cohesion Policy is high as well as in all clusters, nevertheless, people in this group has a higher chance than Cluster 4 or 3 to be aware of any European policy.

“**Confident Europeans**” (Clusters 3), who adhere to the values and identify themselves but consider the action of the European Union to be less effective, and “**Wary pro-Europe**” (Cluster 4), who consider the European Union to be effective, they adhere to values and identify themselves but see a lot of corruption, are similar in many regards: these include people with a high probability to strongly identify with Europe, even more than how much they identify with their countries, which think the EU membership is a good thing and that they strongly agree with the values of solidarity represented by the cohesion policy.

They have a somewhat different view regarding the effectiveness of EU because people in Cluster 3, Confident Europeans, judge the action of EU in solving problems not so effective and less effective than national government action more likely than people in Cluster 4. On the contrary, individuals in Cluster 3, Confident Europeans, trust the EU institutions more than people in Cluster 4, Wary pro-Europe, because they more likely perceive a lower level of corruption in EU, even respect to their national governing institutions. The Clusters 2, 5 and 6 include people with, in comparison with the others Clusters, a weaker identification with Europe: especially in Cluster 2, **EU Deniers**, is null the probability that people strongly identify with Europe, and is 46%, the highest among all clusters, the probability that people do not identify at all with EU. Generally, in each of these three clusters people identify stronger with their own countries. Cluster 2, EU Deniers, emerges as the group characterized by the mostly negative attitude toward many aspects considered by the indicators: they consider (with probability 0.9) the EU is not effective in solving problems of their region and less effective than national government, they think (with probability 0.86) that corruption is widespread in EU institutions, as like as in national institution and even more, and they think that the EU membership is a bad thing (with probability 0.74). Nevertheless, the majority still agree with the policy of sustaining the poorest regions but the proportion of individuals who disagree with this policy is the highest across clusters (36%). Cluster 5 and 6 have in common with Cluster 2 the stronger identification with their own countries than with Europe but the identification with Europe is somewhat stronger and people less likely evaluating the EU membership as a bad thing (they are equally divided). For the majority they agree with the policy of supporting the poorest region. Instead, they show very different attitude regarding two main aspects: “**Disaffected Europeans**” (Cluster 5, cons a not effective EU) negatively evaluate the effectiveness of EU in solving problems in their region (with probability about 0.80) and consider national government as well as or more effective; on the contrary people in Cluster 6 positively evaluate the effectiveness of EU institutions even in comparison to the national institutions.

Regarding the level of corruption, the situation is reversed: in Cluster 5, the probability of considering high the level of corruption in EU institution is zero, while in Cluster 6, **Wary cons Europeans** (cons a corrupted EU) more people perceive higher level of corruption (with probability 0.50) As for cohesion policy, Clusters 2, and Cluster 5 have the largest proportion of people who do not agree with the current financial support policy by EU. Measuring how

many times people voted in the last two elections has the last association with the latent variable  $X$  and so it does not discriminate too much among the clusters; anyway, Cluster 2 and Cluster 5 have the highest probability (40%) of not voting both the times, while this probability is lower in Cluster 3 and 4.

### 4.3 Identifying groups of regions

As for individual clusters, we can obtain a global synthesis of the characteristics also for the latent classes of  $W_j$  from the profile Table 4. The first row shows the size of the classes at regional level  $\hat{P}(W_j = m)$ . Since model specification contains group-level covariates, the probabilities  $\hat{P}(W_j = m)$  are computed aggregating the model probabilities  $\hat{P}(W_j = m | \mathbf{Z}_j^g)$  over the values of group-level covariates. These probabilities are called *prior or model probabilities*.

Moreover, Table 4 shows, for each category of the indicators, the group-region specific probabilities  $P(Y_{ijk} = s_k | W_j = m)$  given the latent class.

Looking at these probabilities, the classes 1 and 4 can be identified both as group of regions which, among all the regions take into account, have the weaker level of identification with Europe, especially in comparison with country identification, and quite critics against the EU institutions. They have quite similar profile along many dimensions: people think that the EU is not very effective in solving problems of the region (with probability of 0.62), and that corruption in EU institutions is widespread. Although the probability to agree with the EU policy of supporting the poorest region is high (77%), this value is lower than that observed in Group 2 and Group 3. Possibly, Group 4 “**Low EU identification – Skeptical**” is characterized by the worst attitude of people toward EU membership of their countries: more than half of the people in regions of Group 4 consider, more than peoples in other groups, the EU membership a bad thing, even more than Group 1 (where the same probability is 0.44). Another difference between the two groups regards the awareness of the Cohesion Policy: the chance that people do not know any EU policy is 27% in Group 1 “**Lower EU identification**” while it is 14% in Group 4, with a greater proportion of people knowing cohesion/regional policies financed by EU.

On the contrary, Group 2 “**High EU identification**” can be labelled as the group of regions with high level of identification with Europe and where people have a higher level of trust and appreciation for EU institution: this group has, in comparison with the other groups, the highest probability of strongly identifying with Europe (0.62), and most people agree to support poorest regions (with probability 0.83). People in these regions have a high chance to consider the EU membership a good thing (with probability 0.71) and they consider EU institutions effective in solving problems with probability greater than 60%; moreover, there is a larger proportion of people that think EU is more effective than national government (probability about 0.3). In addition, this group have the lowest probability to perceive high level of corruption in EU.

**Table 4.** Profile table of the latent variable at regional level  $W_j$ : size class  $\hat{P}(W_j = m | Z_j^g)$  and class specific marginal probabilities  $\hat{P}(Y_{ijk} = s_k | W_j = m)$

	<b>Group 1</b> <i>Lower EU identification</i>	<b>Group 2</b> <i>High EU identification</i>	<b>Group 3</b> <i>Medium-high EU identification – Critics</i>	<b>Group 4</b> <i>Low EU identification – Skeptical</i>
Group Size	0.386	0.322	0.151	0.142
<b>Indicators</b>				
<b><i>How strongly identify with Europe</i></b>				
<i>Not much strongly</i>	0.195	0.119	0.130	0.179
<i>Somewhat</i>	0.360	0.261	0.282	0.318
<i>Strongly</i>	0.445	0.620	0.588	0.504
<b><i>Europe vs Country identification</i></b>				
<i>Less</i>	0.349	0.221	0.241	0.304
<i>Equal</i>	0.569	0.666	0.658	0.605
<i>More</i>	0.081	0.113	0.101	0.092
<b><i>Effectiveness in solving problems</i></b>				
<i>Not so Effective</i>	0.619	0.381	0.577	0.625
<i>Somewhat effective</i>	0.303	0.449	0.335	0.299
<i>Very effective</i>	0.079	0.171	0.088	0.076
<b><i>EU vs National effectiveness</i></b>				
<i>Less</i>	0.256	0.168	0.249	0.259
<i>Equal</i>	0.608	0.534	0.598	0.610
<i>More</i>	0.136	0.298	0.152	0.131
<b><i>Corruption in EU</i></b>				
<i>Low</i>	0.107	0.161	0.111	0.119
<i>Medium</i>	0.374	0.431	0.381	0.383
<i>High</i>	0.519	0.409	0.508	0.498
<b><i>EU vs National Corruption</i></b>				
<i>Less</i>	0.188	0.358	0.170	0.233
<i>Equal</i>	0.636	0.517	0.657	0.597
<i>More</i>	0.175	0.125	0.174	0.170
<b><i>Vote</i></b>				
<i>Neither</i>	0.333	0.303	0.316	0.325
<i>Once</i>	0.161	0.167	0.160	0.159
<i>Both times</i>	0.486	0.509	0.504	0.496
<i>(d/k-refused)</i>	0.021	0.022	0.020	0.020
<b><i>Support to Cohesion policy</i></b>				
<i>Agree</i>	0.771	0.831	0.800	0.779
<i>Disagree</i>	0.216	0.157	0.187	0.208
<i>d/k</i>	0.013	0.012	0.013	0.013
<b><i>EU membership</i></b>				
<i>Bad thing</i>	0.444	0.291	0.238	0.563
<i>Good Thing</i>	0.556	0.710	0.762	0.437
<b><i>Awareness of Cohesion policy</i></b>				
<i>None</i>	0.266	0.123	0.224	0.140
<i>Only local project</i>	0.085	0.122	0.090	0.122
<i>Cohesion/regional policies</i>	0.649	0.755	0.686	0.738



Finally, the regions in Group 3 “**Medium-high EU identification – Critics**” have a profile in between these two extremes. In particular, in these regions there is a relatively high proportion of people that strongly identify with Europe, as strong as with their own country, and about 80% of people that approve the EU financial support of poorest regions. Moreover, the proportion of people that think the EU membership is a good thing is the highest (probability 0.76). However, the perception of efficacy and corruption of EU institutions is not so good as in the regions of Group 2. Actually, people has 57% chance of answering that EU is not effective in solving region’s problems and 25% probability of responding EU is less effective than national government, as well as 51% chance of answering that corruption in EU institutions is high. Moreover, compared to Group 2, it is higher the probability that people do not know any EU financed policy (0.22), a value similar to Group 1 more than to Group 2.

Regional classification reflects, to some extent, individual typologies, even if in this case differences are less marked. Computing the probability of being in a certain latent class of  $X_{ij}$  for each level  $m$  of  $W_j$ ,  $P(X_{ij} = t | W_j = m)$ , we can quantify the influence of the latent classes at individual level across latent classes at regional level.

Table 5 shows the relationship between the individual and the regional classes. The relative size of individual clusters within a group of regions can reveal different structures for the latent variable “identification” across regions, depending on the effect of latent variable at regional level. In this case, there is some overlap in the composition by clusters of each regions’ group. Consistent with the previous analysis, Group 2 is composed for about 50% by Clusters 3 and 4, hence those individuals with a high level of identification and a positive evaluation of EU institution on both dimension. Group 3 is composed by 40% of individual classified in Cluster 1, and in minor proportion by Cluster 4 and 2. Group 1 and Group 4 have quite similar profile, since they are composed for the majority by the two Clusters 1 and 2 with a little bit higher proportion of Cluster 1 in Group 4; the difference is the presence of a certain proportion of Cluster 5 in Group 1 and of Cluster 3 in Group 4. We note the presence of a certain proportion of Cluster 1 in each group, which is a cluster composed by people that strongly identify with Europe but that do not trust European institutions.

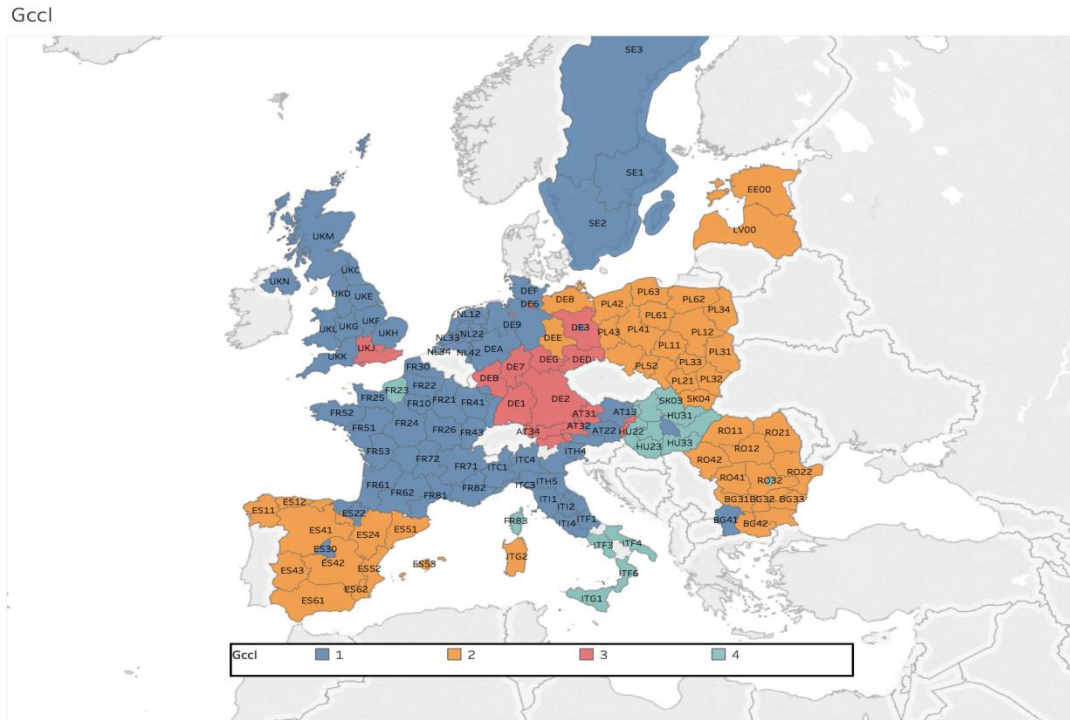
**Table 5.** Cross-tabulation of the probability of being in each latent class of  $X_{ij}$  for each level of  $W_j$ :  $\hat{P}(X_{ij} = t | W_j = m)$

	Group 1	Group 2	Group 3	Group 4	Overall
<b>Group class Size</b>	0.386	0.322	0.151	0.142	
<b>Clusters</b>					
<i>Cluster1</i>	<b>0.284</b>	<b>0.184</b>	<b>0.409</b>	<b>0.319</b>	0.273
<i>Cluster2</i>	<b>0.285</b>	0.106	<b>0.165</b>	<b>0.267</b>	0.197
<i>Cluster3</i>	0.087	<b>0.290</b>	0.091	<b>0.165</b>	0.175
<i>Cluster4</i>	0.108	<b>0.216</b>	<b>0.141</b>	0.088	0.150
<i>Cluster5</i>	<b>0.154</b>	0.060	0.116	0.111	0.106
<i>Cluster6</i>	0.082	0.145	0.078	0.052	0.100

#### 4.4 Mapping the level of identification with EU across European regions

Figure 1 shows the regions' classification across the four groups using a modal assignment based on prior marginal probabilities (the region is assigned to the group corresponding to the highest prior probability).

**Figure 1.** Regions classification in latent variable group  $W_j$  based on prior probabilities



- Group 1 – Lower EU identification
- Group 2 – High EU identification
- Group 3 – Medium-high EU identification – Critics
- Group 4 – Low EU identification – Skeptical

The largest group is Group 1 “Lower EU identification” that comprises most of the regions in the northern countries of Europe: Sweden, almost all regions in UK, except for the London and South-East regions, and North Ireland, Netherlands, almost all regions from France and the northern regions of Germany. Belong to this group also the regions from the North and the Centre of Italy and some Austrian regions. The rest of Italian regions (located in the South) are classified to belong to Group 4 “Low EU identification – Skeptical”, where people are possibly even more critics against EU than people in Group 1. To this group belong also the regions in Hungary and some from Slovakia. On the contrary, all of the Polish regions, the regions from Romania, Estonia and Latvia, and almost all regions from Bulgaria belong to Group 2 “High EU identification”: people identify strongly with Europe, believe the membership of their country to EU is a good thing, and trust EU institutions. Almost all the Spanish regions belong to the same group too, except for three of them. In Group 3 “Medium-high EU identification – Critics” are classified the central regions from Germany and some regions from Austria, plus the regions of London and South-East in UK. This group is characterized by a high level of identification with Europe, nevertheless they evaluate not much positively the effectiveness and the level of corruption of European institutions.

The assignment of regions to a specific level-2 group is better evaluated by considering the *posterior membership probability*, which represents the probability for a given region  $j$  of belonging to latent cluster  $m$  of  $W$  given the responses of all its individuals  $Y_j$  and their covariates  $\mathbf{Z}_j$ :

$$\hat{P}(W_j = m | Y_j = s, \mathbf{Z}_j) = \frac{\hat{P}(W_j = m | \mathbf{Z}_{ij}^g) \prod_{i=1}^{n_j} \hat{P}(Y_{ij} | W_j = m, \mathbf{Z}_{ij})}{\hat{P}(Y_j = s | \mathbf{Z}_j)}$$

We can obtain a classification table that compares the results of the classification through modal assignment based on posterior probabilities (Table 6). The classification error is 0.024, which means that only few regions have similar posterior probabilities across some groups, in particular for regions classified in Group 1 and 4.

**Table 6.** Classification table of regional latent variable  $W_j$  based on group-level posterior membership probabilities  $\hat{P}(W_j = m | Y_j = s, \mathbf{Z}_j)$

Posterior group membership probability	Modal classification				
	1	2	3	4	Total
1	57.38	0.29	0.32	0.36	58.35
2	0.02	48.67	0.00	0.00	48.70
3	0.95	0.00	21.68	0.00	22.63
4	1.65	0.04	0.00	19.64	21.33
Total	60	49	22	20	151

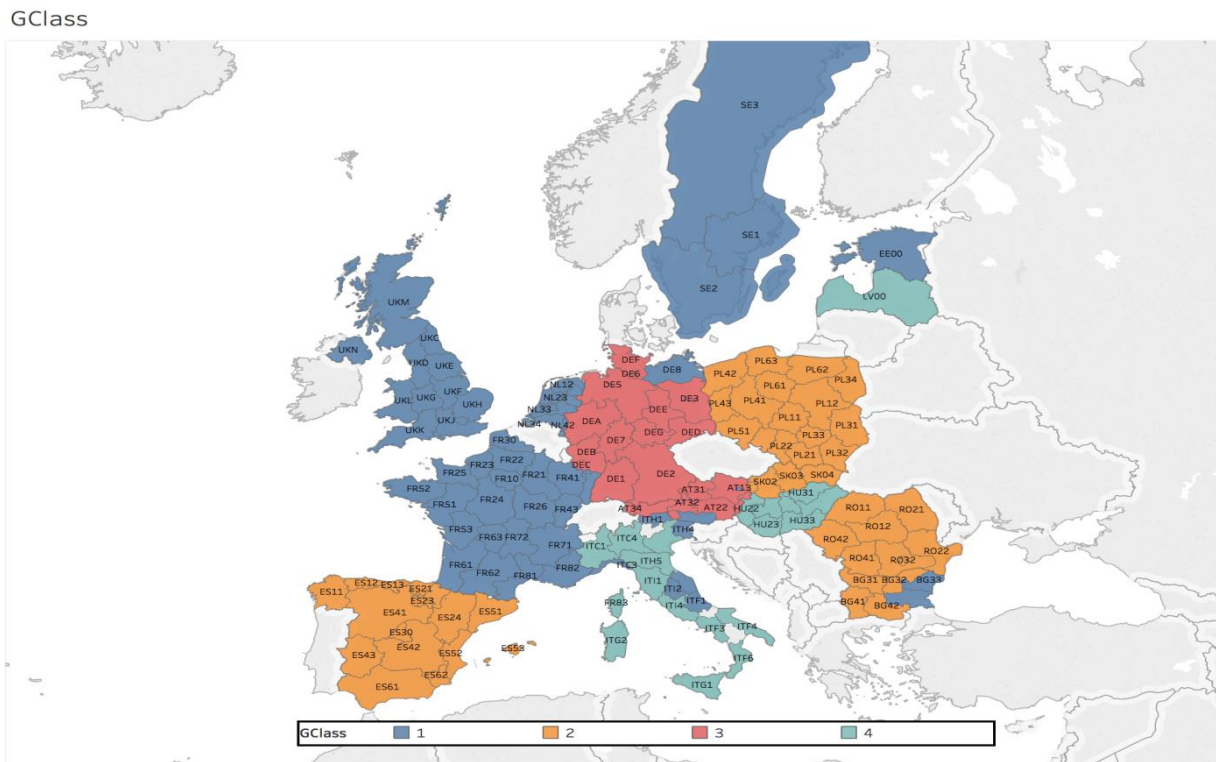
Figure 2 shows the assignment of the regions to the different groups based on posterior probabilities. The comparison between the two probabilities sets (prior and posterior) highlights to what extent the individual responses can change the probability to belong to a specific group. In fact, while the prior probabilities are obtained according to the group level covariates only (i.e. are model's parameters), the posterior probabilities are conditioned both on the individual responses and covariates patterns. Discrepancies between the two set of probabilities highlight that the group level covariates are useful to describe the socio-economic context. The inclusion of individual responses adds the relevant information which allows for a better prediction of the group membership for the regions concerned. In particular, 33 out of 151 regions are classified differently after accounting for the individual responses. These are the regions in the North of Italy, some German and Austrian regions, Latvia and Estonia, plus others single regions in different countries (three regions in each country of Spain, Slovakia, and Bulgaria; two regions from UK, one each from Hungary, France, and Romania). Group-level probabilities and respective modal group assignment corresponding to the two sets of probabilities are reported in Table A.3 in Appendix (the thirty-three regions that change classification are highlighted in grey). In many cases, the group prior probabilities are close each other. For example, in some regions the membership probabilities are almost equally divided by two specific groups: in this case the change of group membership regards the two groups with the highest prior probabilities. The regions Niedersachsen (D9) and Schleswig-Holstein (DEF) have a prior probability close to 0.50 to be classified in Group 1 and Group 3, as like Steiermark (AT22), Ugozapaden (BG41),

Trentino-Alto Adige (ITH1), Western Slovakia (SK02) and Central Slovakia (SK03), and South-East England (UKJ). In other cases, instead, the region has significant prior probability to be classified in every group: Sachsen-Anhalt (DEE) has probabilities between 0.22-0.29 to belong to each group from 1 to 4; Mecklenburg-Vorpommern (DE8), Comunidad Foral de Navarra (ES22), Central Hungary (HU10), Bratislava Region (SK01), and Estonia all have three membership prior probabilities equal at least 20% or more. Generally, change in classification happens between the two groups with the highest probabilities.

There are some exceptions. Some regions in Italy (ITC1: Piemonte, ITC4: Lombardia, ITH3: Veneto, ITH5: Emilia-Romagna) have very high prior probability to belong to Group 1 and are classified to Group 4 by posterior probability; however, the profile of two groups are quite similar in many respects. Latvija has a high prior probability to be classified in Group 2 (0.80) and probability 0.19 to be classified in Group 4, which become equal to 1 if we consider posterior probability. The region Bucuresti-Ilfov (RO32) in Romania, on the contrary has probability 0.8 to belong to Group 4 and probability 0.15 to belong to Group 2 but considering the pattern of responses it is assigned to Group 2 with certainty. Two regions from Bulgaria (BG33: Severoiztochen and BG34: Ugoiztochen) are classified in Group 2 with very high prior probability but in Group 1 by posterior probability, while one region in Germany (DE3: Berlin) passed from Group 1 (with prior probability 0.80) to Group 3 (with posterior probability 1). For most of these cases, posterior probabilities are either zero or one, so we can conclude that the context covariates considered in the analysis do not allow to predict group membership very well, at least for these cases.

In comparison with Figure 1, we see now that all the UK regions and the North Ireland, all the French regions, Netherland, Sweden and Estonia are more likely to be classified in Group 1, which together with regions in Group 4 are characterized by a lower level of identification and the most critics against EU institutions. Now most of the German regions, except Mecklenburg-Vorpommern, and the regions from Austria are classified in Group 3, which is a group that identify with EU even if they are critics about some aspects. Most of the regions in countries of Eastern Europe, except the regions from Hungary, are classified in Group 2: they have the highest level of identification with Europe and trust EU on effectiveness and corruption. In this group are even all the Spanish regions. Note that Estonia and Latvija are now classified quite differently: the former is classified in Group 1 and the latter in Group 4 that are regions that identify the lowest (previously, they were both classified in Group 2). Group 4 now comprises the Northern Italian regions as well, which change classification from Group 1 to Group 4, and it includes still all the regions from Hungary, and Corsica.

**Figure 2.** Regions classification in latent variable group  $W_j$  based on posterior probabilities



- Group 1 – Lower EU identification
- Group 2 – High EU identification
- Group 3 – Medium-high EU identification – Critics
- Group 4 – Low EU identification – Skeptical

#### 4.5 Ranking of the case study regions

The nine case-study regions of the PERCEIVE project have been selected according to their ability to represent the complex and heterogeneous reality of EU Cohesion Policy performance and its multidimensional determinants in terms of socio-economic, political and demographic development. The sample is balanced between regions targeted for the “Competitiveness and Employment” and for the “Convergence” objectives, with Emilia-Romagna and Calabria accounting for the within-country heterogeneity in European Structural Funds support in Italy. Extremadura is the sole Spanish region that stayed under the “Convergence” objective in the programming period 2014-2020, whereas Burgenland (Austria) was targeted as a former “Convergence” objective region but shifted to “Phasing out” in the 2007-2013 programme.

Great variability in the economic performance can be observed (Table 7) among regions<sup>4</sup>. In terms of GDP per capita the three “Competitiveness” Objective regions, Emilia-Romagna, Norra Mellansverige and Essex, are in the top 25% percentile of the distribution of the regions in EU. The regions covered by the “Convergence” Objective are amongst the poorest in Europe. The Romanian and Polish regions’ GDP per capita in 2014 were below the 50% of

<sup>4</sup> A detailed socio-economic analysis of the case-study regions is provided in D1.1 “Report on Regional Case-Studies”.



the EU average, Calabria and Extremadura's were below the 60% of the EU average, while the GDP of the Burgenland was very close to the European average.

The unemployment rate mirrors the Convergence/Competitiveness distinction, too: if the Competitiveness regions showed an unemployment rate below the 8%, Convergence regions' unemployment rate ranges from 9% in Dolnoslaskie up to 29% in Extremadura.

**Table 7.** Case study regions, socio-economic characteristics

Region	Objective	GDP*	Unemployment**	EQI	Absorption rate***	Structural Funds per capita****
Sud-Est (RO)	CONV	25	10,00	13,37	32,14%	934,09
Dolnoslaskie (PL)	CONV	43	9,10	35,49	64,70%	1.741,62
Warminsko-Mazurskie (PL)	CONV	28	9,60	43,99	66,33%	2.058,60
Extremadura (ES)	CONV	56	29,30	54,35	116,02%	2.359,96
Burgenland (AT)	PH-OUT	97	4,70	68,14	77,37%	625,02
Emilia-Romagna (IT)	COMP	118	8,20	31,77	44,44%	107,78
Calabria (IT)	CONV	59	23,10	17,85	50,24%	1.491,24
Essex (UK)	COMP	117	4,50	65,54	59,84%	67,68
Norra Mellansverige (SE)	COMP	139	7,20	74,24	79,09%	436,51

\* Euro per inhabitant in percentage of the EU average, 2014

\*\* Unemployment rate: population 20-64 years, 2014

\*\*\* Absorption rate of SF expenditures: ratio of SF expenditures up to 2013 to the SF allocation in the 2007-2013 period

\*\*\*\* Total expenditures financed by Structural Fund in years 2007-13 divided by the average population in a region in the period 2007-13

With respect to quality of institutions, regions from Austria, United Kingdom and Sweden hold the top positions of Europe, but the rank can vary when considering different pillars of the index. Although the region Emilia-Romagna is one of the best performers in Italy, it ranks quite low compared to other European regions due to citizens' perception of corruption and personal experiences with bribery in the public sector. Among the sole Convergence Objective regions, the situation is much more differentiated. Calabria has one of the lowest values of the index in Europe, and performs badly in comparison to other Italian regions, too. Romania, among the countries with the lowest performance, has a high level of perceived corruption especially in public services, and the Sud-Est (RO) region does not differ from Romania's national case. Poland's institutional quality is below the EU average: Dolnośląskie and Warminsko-Mazurskie have both negative score but they have the lowest and the best scores, respectively, of all the Polish regions, with the former exhibiting a poor quality of the public health care system and police force. Somewhat different is the situation of Extremadura (ES), which ranks very similar to the European average and slightly higher than the whole Spain, even if they both experienced a decline of the quality of institutions index.

In 2013 the Structural Funds absorption rate was very differentiated across regions, ranging from 116% in Extremadura to 32% in Sud-Est (RO). The absorption rate is distributed among regions regardless of their policy objective and the amount of resources available. Looking at the Structural Funds per capita, the most benefitted region was Extremadura with €2,359, followed by the two Polish regions and Calabria with €1,491. The remaining regions, with Structural Funds per capita below €1,000, range from €934 in Sud-Est to €67 in Essex.

The regions' high variability in terms of macroeconomic context, institutional quality and Cohesion Policy's financial allocation reflects into the highly differentiated picture that emerges when looking at the regions' levels of identification with the EU (Table 8).

**Table 8.** Ranking of the case study regions according to their level of identification

Region	Objective	Group	Name
Sud-Est (RO)	CONV	2	High EU identification
Dolnoslaskie (PL)	CONV	2	High EU identification
Warminsko-Mazurskie (PL)	CONV	2	High EU identification
Extremadura (ES)	CONV	2	High EU identification
Burgenland (AT)	PH-OUT	3	Medium-high EU identification – Critics
Emilia-Romagna (IT)	COMP	4	Low EU identification – Skeptical
Calabria (IT)	CONV	4	Low EU identification – Skeptical
Essex (UK)	COMP	1	Lower EU identification
Norra Mellansverige (SE)	COMP	1	Lower EU identification

The groups 2 and 3 show higher levels of EU identification and sense of attachment to the European project.

Four out of nine case-study regions are classified within Group 2 “High EU identification”, with three of them being part of new accession countries Poland and Romania. These regions feature a higher degree of identification with Europe, and their citizens think that EU institutions are effective in facing their region's problems. Thirty percent of them also think that corruption is more spread in their national institutions than in the European Union's.

These four regions share a lower than the EU average level of quality of institutions, and are among the highest beneficiaries of the Structural Funds, along with Italy's Calabria.

The only case-study region that belongs to Group 3 “Medium-high Identification – Critics” is Burgenland, who detaches from Group 2 because of its citizens' skepticism towards EU institutions. Regions in Group 3 do show identification with Europe and support for the European project, yet their citizens think that EU institutions are corrupt and not effective in facing their regions' needs.

Essex and Norra Mellansverige belong to the group that includes the highest number of European regions: Group 1 “Lower EU identification”. These two regions are among the richest of the sample and feature a high level of institutional quality, yet their citizens hardly identify themselves with Europe, nor do they consider the EU capable of tackling the problems and needs faced by their region. They also think that EU institutions are corrupt, more than their national institutions, and even though they sort of know and appreciate Cohesion Policy they do not think that the EU membership brings benefits to their country.

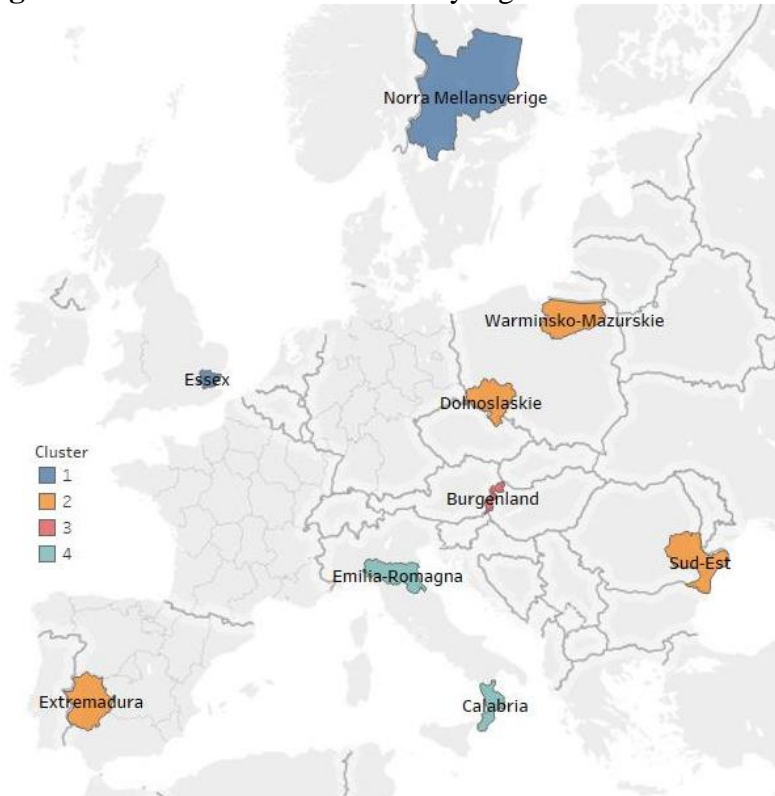
Among the case-study regions of the PERCEIVE project, both Italian regions, Emilia-Romagna and Calabria, are classified into Cluster 4 which represent a particular case. Despite a higher knowledge of Cohesion Policy, their citizens are negative about EU membership, consider the EU corrupt as much as their national institutions, and not effective in solving their problems.

Regardless of the differences of the social and economic contexts, the citizens of Emilia-Romagna and Calabria share a common mistrust towards institutions, both national and European, and consider Italy's adhesion to the EU a bad thing for their country. However, considering the posterior membership probabilities, conditioned on individual responses and covariate patterns (see section 4.4), the differences in economic and social contexts become less relevant: Emilia-Romagna, initially placed in Group 1 (where citizens have shown a very low identification with the European Union), after considering individual responses, enters the same group as Calabria (Group 4), also characterized, as we said, by mistrust of all institutions.

Our results are consistent with the findings from D2.3 (Royuela, 2018). In both analyses the Eastern European regions of Romania (Sud Est) and Poland (Dolnoslaskie and Warminsko-mazurskie) display the highest values of the Composite Index of European Identity, followed

by Extremadura (Spain), while the lowest values of the composite index are reported for Norra Mellansverige (Sweden) and for Essex, which are here classified in Group 1 “Lower EU identification”. If Royuela (2018) finds average levels of the index in the regions of Burgenland (Austria), Emilia-Romagna and Calabria (Italy), in our analysis these regions are placed in two intermediate groups (respectively Group 3 “Medium-high EU identification” Discouraged and Group 4 “Low EU identification – Skeptical”). These two groups have in common the critical attitude toward the effectiveness and corruption of EU. However, citizens of regions in Group 3 have a positive opinion about their country’s EU membership and show strong identification with Europe, while citizens of regions in Group 4 do not.

**Figure 3.** Classification of case-study regions



The three dimensions of “Awareness” (in terms of the EU Cohesion Policy and its local benefit), “Evaluation” (in terms of perceived effectiveness of political institutions and benefits of EU membership) and “Attachment” (participation in the election and support to the redistributive role of the EU), play a fundamental role in determining a region's group membership. The awareness of EU Cohesion Policy is highest among Sud-Est, Dolnoslaskie, Warminsko-Mazurskie and Extremadura, regions of Group 2 “High EU identification” and lowest among regions in Group 1 “Lower EU identification” (Essex and Norra Mellansverige).

“Attachment” follows an analogue pattern: citizens of the regions in Group 2 are those who are most likely to have voted in both last elections for the European parliament and show a high level of support for the EU redistributive policy, followed by the citizens of the regions of Group 3 (Burgenland), 4 (Emilia-Romagna and Calabria) and 1.

The “Evaluation” dimension, on the other hand, is the least coherent of the three. Group 2 have the highest probability of finding citizens who consider the work of the EU effective and who perceive a low level of corruption in the European institutions but, in the other

groups, this dimension is more ambiguous. Irrespective of their level of identification with Europe and the European project, the citizens in regions belonging to Groups 1, 3 and 4 consider the European institutions as corrupt and ineffective as much as their national peers.

#### 4.6 Effects of covariates

Membership to individual latent classes is related to individual demographic or socio-economic characteristics. The probability that an individual belongs to a particular latent class has been modelled to depend on age of individuals (considered as continuous variable), the level of education, and the income level. The effects of covariates depend on logit parameters. However, in order to facilitate the interpretation, we present the class membership probabilities given each level of each covariate,  $P(X = t | \mathbf{Z}_p = z)$ , computed aggregating over the categories of the other covariates, and over the latent variable at group level  $W_j$  (Table 9).

For each row in correspondence to a specific range of values for a covariate the values reported are the probabilities to be classified in each cluster, which can be compared to the overall distribution reported in the first row (the latent class sizes - unconditional probabilities). The variable age is introduced in the model as a continuous variable, but the probabilities are calculated for age classes (automatically determined dividing the sample values in five groups). Comparing, the distribution for each age class, to the marginal distribution by cluster, we notice that younger individuals are somewhat underrepresented in Cluster 1 and overrepresented in Cluster 3 (which is the cluster most favourable to Europe and EU) and Clusters 5 and 6, while older individuals (especially with more than 65 years) are overrepresented in Cluster 1 and underrepresented in Clusters 3, 5 and 6. Generally, however, it seems that age has not too much influence on identification.

Regarding education, clearer patterns emerge: individuals with a tertiary education have a higher probability than average to be allocated in Cluster 3 and 4 (higher level of identification and positive attitude toward EU institutions) and even to Cluster 1 (strongly identify with Europe but critics with EU institution), while a lower than average probability to be classified in Cluster 2, 5 and 6 (all of them characterised by a lower level of identification). On the contrary, people with the lowest level of education are much more underrepresented in clusters 3 and 4 and overrepresented in Cluster 2, which is the cluster of citizens that do not identify with Europe and the most critics against the EU.

A similar pattern emerges for income level: generally, individuals with high income are most likely than average to be included in Cluster 3 and 1 and less likely in Cluster 2, while individuals with a low level of income are more likely to be classified in Cluster 2 and less in Clusters 3 and 4.

We report also other individual characteristics that are not considered in the final model: although they do not concur directly to determine the classification they can be used to describe the clusters, as in standard “multivariate cluster analysis”.

As for the composition by sex, we can see that there is no a clear divide by gender. About occupational status, we notice that unemployed are over-represented in Cluster 2 and under-represented in Cluster 3, 4 and 1. Students and trainees are more represented in Clusters 3, 4 and 5, and less in Clusters 2 and 1.

Finally, we consider whether people are satisfied with the economic situation in the region at the time of the survey. Indeed, people unsatisfied with the economic situation are over-represented in Cluster 2 and underrepresented in Cluster 4 and 1; the contrary happens for satisfied people.

In conclusion, despite it is possible to identify some influence of individual characteristics on the level of identification and citizens' attitudes toward the EU, in general we cannot recognise a strong dependence of identification on demographic or socio-economic characteristics.

**Table 9.** Conditional probabilities of  $X_{ij}$  for individual level covariates:  $P(X = t | Z_p = z)$ .

	<b>Cluster1</b> <i>Disappointed pro- Europe</i>	<b>Cluster2</b> <i>EU Deniers</i>	<b>Cluster3</b> <i>Confident Europeans</i>	<b>Cluster4</b> <i>Wary pro- Europe</i>	<b>Cluster5</b> <i>Disaffected Europeans</i>	<b>Cluster6</b> <i>Wary Cons- Europe</i>
<b>Overall</b>	0.273	0.197	0.175	0.150	0.106	0.100
<b>Age (years)</b>						
18-33	0.231	0.178	0.189	0.150	0.127	0.125
33-44	0.262	0.190	0.195	0.139	0.107	0.107
45-55	0.288	0.208	0.166	0.143	0.103	0.092
56-65	0.278	0.209	0.166	0.159	0.101	0.087
More than 65	0.303	0.202	0.161	0.156	0.091	0.088
<b>Education</b>						
Primary and lower secondary	0.261	0.245	0.116	0.131	0.124	0.123
High secondary	0.251	0.213	0.177	0.146	0.105	0.109
Degree and PhD	0.297	0.161	0.202	0.161	0.098	0.081
<b>Income</b>						
Low	0.245	0.217	0.156	0.142	0.113	0.127
Medium	0.249	0.215	0.166	0.160	0.110	0.101
High	0.312	0.170	0.192	0.149	0.096	0.081
DK/R	0.285	0.180	0.209	0.141	0.105	0.080
<b>Not Included in the Model</b>						
<b>Gender</b>						
Male	0.263	0.2049	0.1732	0.1517	0.1069	0.1002
Female	0.2822	0.1896	0.1772	0.1473	0.1042	0.0995
<b>Occupation</b>						
Employed	0.2709	0.1912	0.1827	0.1457	0.1071	0.1023
Unemployed	0.2314	0.2544	0.1346	0.1386	0.1147	0.1263
Housewife, Pensioner, Retired, Other	0.2878	0.2046	0.1646	0.1555	0.097	0.0905
Student, Trainee	0.2397	0.1352	0.2159	0.1695	0.1383	0.1014
<b>Satisfied with economic situation</b>						
Not satisfied	0.2375	0.238	0.1745	0.1327	0.1068	0.1105
Satisfied	0.2997	0.1658	0.1758	0.1623	0.1046	0.0917



We can replicate the analysis considering also the membership probabilities (for individual level and group-level latent variables) by specific values of the context covariates at regional level (Table 10 and Table 11 respectively).

The wealth of regions where people live have some influence on the probability to be classified in different clusters. Cluster 1 is over-represented in regions with high GDP and low unemployment. Cluster 3 and 4 are overrepresented in poorest regions and underrepresented in regions with high levels of GDP and low level of unemployment. Cluster 2, the most adverse to EU project, are underrepresented in the poorest regions and overrepresented in richest regions while the level of unemployment in the region does not seem to influence too much. Cluster 6 is more likely present in very poor regions and high unemployment, on the contrary of Cluster 5 that is underrepresented in the same regions. Looking at the effect of absorption rate, that is a proxy of the regions' effectiveness and efficiency in programming and implementing Cohesion Policy, not a clear pattern can be observed, apart for Clusters 1 and 3. People living in regions with high values of the rate are most likely assigned to Cluster 1 and somewhat to Clusters 4 and 6, while people living in regions with low values (less efficient) are more likely assigned to Cluster 3.

The amount of Structural Fund received in the region seems to be more influencing: people living in regions receiving more funding (typically regions targeted as "Convergence Objective") are more likely classified in Clusters 3 and 4 (typically with strong identification and positive attitude toward EU); on the contrary, people living in regions where EU funding is low are more likely to be in Clusters 1 (they strongly identify but they do not trust EU), in Cluster 2 (do not identify and do not trust EU), and 5 (they identify more with their country, and do not evaluate much the effectiveness of EU in solving problems).

The last variable reported in the table regards the Quality of Institution Index. It is not included in the final model as an active covariate (i.e. it does not affect the membership probabilities) because it was highly correlated with GDP and SF expenses; at the same time, we already account for perceived corruption among the manifest variables. However, we can describe each cluster by this variable too in order to check the correspondence with this indicator at regional level. In correspondence to high values of the index (high institutional quality in the region), individuals have a higher chance to be classified in Cluster 1, 2 and 5: although different for the level of identifications, these clusters have in common people perception of high level of corruption in EU institutions and more trust in their national governments. On the contrary, in correspondence of low and intermediate levels of this indicator, people more likely are classified in Clusters 3 and 4, both with high level of identification, even if individuals in Cluster 4 are critics about corruption in EU, and have somewhat higher chance to be classified in Cluster 6 too that have a lower level of identification with EU but they trust EU institutions regarding corruption.

**Table 10.** Conditional probabilities of  $X_{ij}$  for contextual level covariates:  $P(X = t | Z_p = z)$

	<b>Cluster1</b> <i>Disappointend pro- Europe</i>	<b>Cluster2</b> <i>EU Deniers</i>	<b>Cluster3</b> <i>Confident Europeans</i>	<b>Cluster4</b> <i>Wary pro- Europe</i>	<b>Cluster5</b> <i>Disaffected Europeans</i>	<b>Cluster6</b> <i>Wary Cons- Europe</i>
Overall	0.2727	0.1972	0.1752	0.1495	0.1056	0.0999
<b>Covariates</b>						
<b>GDP</b>						
0-37	0.194	0.120	0.302	0.169	0.074	0.141
37-59	0.248	0.182	0.191	0.174	0.101	0.105
59-97	0.248	0.210	0.175	0.149	0.103	0.116
97-118	0.302	0.273	0.118	0.121	0.122	0.065
More than 118	0.375	0.203	0.088	0.133	0.129	0.073
<b>Unemployment rate</b>						
0-5.5	0.350	0.197	0.126	0.136	0.112	0.079
5.5-8	0.285	0.229	0.134	0.127	0.136	0.090
8-9.5	0.281	0.187	0.191	0.156	0.097	0.088
9.5-14	0.240	0.193	0.219	0.130	0.102	0.117
more than 14	0.205	0.179	0.211	0.199	0.080	0.127
<b>Absorption rate</b>						
0-0.5	0.224	0.188	0.234	0.133	0.097	0.124
0.5-0.6	0.272	0.221	0.164	0.166	0.100	0.077
0.6-0.65	0.315	0.195	0.175	0.135	0.103	0.076
0.65-0.78	0.316	0.189	0.146	0.145	0.113	0.091
more than 0.78	0.232	0.194	0.158	0.170	0.115	0.132
<b>SF per capita</b>						
0-108	0.314	0.232	0.120	0.125	0.126	0.083
108-215	0.310	0.248	0.118	0.126	0.121	0.077
215-935	0.239	0.160	0.228	0.151	0.100	0.123
935-2059	0.267	0.166	0.203	0.161	0.083	0.119
More than 2059	0.235	0.180	0.207	0.183	0.097	0.099
<i>Not included in the model</i>						
<b>Quality of institution Index</b>						
0-36	0.228	0.192	0.243	0.126	0.092	0.119
36-42	0.250	0.161	0.232	0.187	0.078	0.092
42-56	0.216	0.181	0.204	0.171	0.099	0.128
56-66	0.318	0.228	0.119	0.132	0.125	0.078
More than 66	0.353	0.223	0.077	0.131	0.133	0.083

The probability that a region belongs to a group identified by a particular latent class at level-2 is modelled to depend on its socio-economic characteristics as well. Table 11 shows the group-level class membership probabilities by the levels of each contextual covariate.

The influence of the level of GDP on group-level membership probabilities is quite evident: regions with higher level of GDP than the European average are overrepresented in Group 1 “Lower EU identification” and Group 3 “Medium-high EU identification – Critics”; on the contrary, regions with lower value of GDP than EU average are more likely to be classified in Group 2 “High EU identification”; in Group 4 “Low EU identification – Skeptical” are more likely classified the regions with a level of GDP below 50-60% of the EU average.

Regarding the unemployment rate, regions with smaller rates are overrepresented in Group 3 (Identify but Critics) and underrepresented in Group 2 (High identification) and Group 4 (Weak identification and skeptical), in particular regions with very high unemployment rate are more likely classified in Group 2 (High Identification and Trust in EU). Group 1 (Weak identification) have instead a mixed composition when considering unemployment rate, because there are regions with either low and high unemployment rate, the only exception is the underrepresentation in this group of regions with very high unemployment rate (on the contrary of Group 4).

As for the influence of the absorption rate, its effect is not so clear cut (as we already notice in the case of individual cluster membership): region with very low values of absorption rate are overrepresented in Group 2 and Group 4 that have very different identification level and trust in EU, and underrepresented in Group 1 and Group 3 (both critics toward EU with regards to effectiveness and corruption, though with a different level of identification). However in group 2 are overrepresented regions with very high absorption rate as well; nevertheless, most of the regions with medium-high values of the absorption rate are more likely classified in Group 1 and Group 3.

Regions that receive more financial funding from EU are most likely classified in Group 2 and Group 4, while regions that receive lower amount of funding from EU are overrepresented in Group 1 and Group 3. This situation reflects what happens in the case of GDP: richest regions are receiving less financial funding from EU hence the groups are characterized by an opposite effect of the two covariates.

Finally, even if it has not a direct influence on regions' membership probabilities, we look at the distribution across groups by different levels of the EQI indicator. In this case as well, the picture emerging characterize Group 1 and Group 3 as composed more likely by regions with high level of institutional quality, while the regions with lower levels of the EQI index are more likely included in Group 2 and Group 4.

**Table 11.** Conditional probabilities of  $W_j$  for contextual covariates:  $\hat{P}(W = m | Z_q^g = z)$

	Group 1	Group 2	Group 3	Group 4
<b>Overall</b>	0.386	0.322	0.151	0.142
<i>Covariates</i>				
<b>GDP</b>				
0-37	0.051	0.739	0.016	0.194
37-59	0.008	0.615	0.061	0.317
59-97	0.361	0.403	0.086	0.151
97-118	0.687	0.069	0.137	0.107
More than 118	0.515	0.003	0.418	0.064
<b>Unemployment rate</b>				
0-5.5	0.387	0.091	0.457	0.065
5.5-8	0.540	0.120	0.233	0.107
8-9.5	0.383	0.428	0.028	0.161
9.5-14	0.404	0.411	0.010	0.174
more than 14	0.103	0.677	0.000	0.220
<b>Absorption rate</b>				
0-0.5	0.198	0.451	0.033	0.318
0.5-0.6	0.421	0.259	0.104	0.216
0.6-0.65	0.448	0.309	0.183	0.060
0.65-0.78	0.433	0.225	0.267	0.076
more than 78	0.390	0.417	0.161	0.032
<b>SF per capita</b>				
0-108	0.461	0.171	0.316	0.052
108-215	0.679	0.116	0.130	0.075
215-935	0.371	0.350	0.131	0.148
935-2059	0.051	0.687	0.080	0.182
More than 2059	0.001	0.443	0.055	0.501
<b>Quality of institution Index</b>				
0-36	0.226	0.537	0.020	0.216
36-42	0.147	0.474	0.039	0.340
42-56	0.164	0.664	0.034	0.138
56-66	0.549	0.175	0.188	0.088
More than 66	0.606	0.009	0.360	0.026

## 5. Concluding remarks

In this deliverable we present the characteristics, specification and results of the novel *IdentEU* probabilistic model. *IdentEU* represents the first attempt to classify EU citizens and regions according to their level of identification with the European project.

*IdentEU* is a Latent Class (LC) model that allows to cluster citizens and regions according to their different patterns of identification with Europe and the European project. In particular, through latent categorical variables we can account for the three main dimensions at the base of the concept of individual identification, i.e., “awareness”, “evaluation” and “attachment” (see section 2.1.2), and for spatial characteristics of European identity and identification discussed in D5.1, D2.2, D2.3. Then, we consider also the influence of individual characteristics and regional context variables. Thus, *IdentEU* contributes to advance our understanding and the assessment of identification and of the factors that influence people's attitudes toward the EU and European project, both at individual and regional level.

The novel *IdentEU* model identifies six clusters of respondent types at the individual level: three clusters include individuals who strongly identify with the European project (Clusters 3, 4, 1), while three clusters include individuals who mostly don't identify with the European project (Clusters 2, 5 and 6). The probability of belonging to the three Clusters 3, 4 and 1 is around 60%. In contrast, the probability of belonging to Clusters 2, 5 e 6, i.e. people who identify less with the European project, is around 40%.

The clusters are heterogeneous with regard to socio-economic and political dimensions, such as the national vs EU identification, the evaluation of the EU membership and its effectiveness, the level of citizens' awareness of the existence of the Cohesion Policy and the agreement on its solidarity value, the trust in EU institutions and the perceived level of corruption.

In particular, at one of the extremes of this classification, we can identify the “**Confident Europeans**” (Cluster 3), who adhere to the EU values, identify themselves as European, and trust the EU institutions because of their lower perceived level of corruption, even if they consider the action of the European Union to be less effective than that of national governments.

At the other extreme, we identify the “**EU Deniers**” (Cluster 2) characterized by the most negative attitude toward the EU, due to the perception of low effectiveness in solving problems at regional or national level, the presence of widespread corruption within the EU institutions (similar or higher than in national institutions), and the negative connotation of the EU membership. In addition, the support to the Cohesion Policy and GDP redistribution to poorest regions in the EU is the lowest among the identified clusters. Nevertheless, Cluster 2 also shows the lowest active participation in the political debate and elections.

Moving to the identification of groups at the regional level, we see a similarity, with regard to the determinants of identification, with the clusters identified at the individual level.

In particular, Group 4 “**Low EU identification – Skeptical**” – and Group 1 “**Lower EU identification**” - are characterized by the citizens' worst attitude toward the EU membership of their countries, while Group 2 “**High EU identification**” represents the group of regions with the highest level of identification with Europe, where people have a higher level of trust and appreciation for EU institution.

It emerges that the UK regions and North Ireland, all the French regions, Netherland, Sweden and Estonia are more likely to be classified in Group 1, which together with regions belonging to Group 4 (which is mostly composed by Italian and Hungarian regions) are characterized by lower levels of identification and higher critical views of EU institutions.

Despite sharing criticism towards institutions, most of the German and Austrian regions strongly identify with EU and belong to Group 3 “**Medium-high EU identification** –

**Critics**". Several of the Eastern European regions belong to Group 2 **"High EU identification"** and show **the highest level of identification with Europe**, trust the EU and considers it more effective and less corrupt than their national governments.

The model results show that the membership into individual latent classes is influenced to some extent by individual demographic or socio-economic characteristics. Indeed, education has a certain influence on citizens' identification with the EU. In contrast, the age dimension has limited influence. In particular, individuals with a tertiary education have a higher probability to be allocated in clusters showing high level of identification and positive attitude toward EU institutions (i.e. Clusters 3, 4 and 1). At the opposite, citizens with the lowest level of education are overrepresented in Cluster 2, which is the cluster that identify weakly with Europe and the most critics against the EU. A similar pattern emerges for income level, in so far individuals with high income are most likely than average to be included in Cluster 3 and 1 and less likely in Cluster 2, while low-income individuals are more likely to be classified in Cluster 2.

In conclusion, we find that, despite being possible to identify some influence of individual characteristics on the level of identification and citizens' attitudes toward the EU, in general we cannot recognise a very strong dependence of identification on demographic or socio-economic characteristics of individuals.

A similar conclusion regards the influence of contextual covariates in predicting the regions' membership to groups with different pattern of identification: 20% of the regions switch their group membership after accounting for the individuals' responses from the survey, whose results create a "new" map of EU identity (section 4.4).

The core of the Northern countries belonging to the EU, with the exception of Estonia (among the recent EU members), shows regions whose citizens identify themselves with the EU project at a lower extent than the others. In fact, all regions in the United Kingdom and Northern Ireland, all French regions, the Netherlands, Sweden and Estonia are more likely to be classified in Group 1, which is characterized by a lower level of identification and the most critical positions towards EU institutions.

Moreover, a relevant number of regions, mostly from Germany and Austria, are classified in a group that identify with the EU, even if they are critics about some aspects (Group 3). In addition, most of Eastern Europe regions (except regions from Hungary) and all Spanish regions have the highest level of identification with Europe and trust the EU for both its effectiveness and absence of corruption (Group 2).

Taking into account the influence of contextual covariates (absorption rate, amount of EU structural funds per capita and level of quality of institutions) we note that: absorption rate is higher in regions with a higher level quality of institution (EQI indicator) but the amount of structural funds per capita received from these regions (generally Competitiveness Objective) is lower. Most regions with these characteristics are classified in Group 1 "Lower EU Identification", therefore adhesion to the EU project is not led by Cohesion Policy.

In the Convergence Objective regions, where the absorption rate is lower as well as the institutional quality, the amount of funds is high, most of these regions are classified in Group 2 "High EU identification" and citizens better perceive the benefits of the Cohesion Policy and identify more with the EU.

However, we see that regions that can be considered similar with respect to contextual characteristics may have quite different level of identification and attitudes toward EU: as Group 2 and Group 4 that have similar characteristics and different pattern of identification, as well as Group 1 and Group 3.



The Northern Italian regions are quite a peculiar case. Regardless of the variety of the social and economic contexts they live in, the citizens of Northern Italian regions share with the southern regions in Italy a common mistrust towards institutions, both national and European, and consider Italy's adhesion to the EU a bad thing for their country (Group 4). In fact, most of them are initially placed in Group 1 together with French and North-European regions, where citizens show a relatively lower identification with the European Union but still considers EU membership in a more favourable manner. However, after considering individual responses, they were classified in Group 4, which is characterized by mistrust towards all institutions.

Our results highlight another crucial aspect: the role of the communication of the Cohesion Policy, whose importance can be indirectly deduced by looking at the regions' change of groups. How are perceptions of EU Cohesion Policy formed at individual level? If our analysis show that these are seemingly independent from the social and economic contexts (sections 4.4 and 4.5.), then there might be a link with the discourse about Cohesion Policy built by the media. This relationship will be further investigated within the PERCEIVE project in D3.4, that will deal with the communication of Cohesion Policy.

The relevance of the three dimensions “Awareness” (in terms of the EU Cohesion Policy and its local benefit), “Evaluation” (in terms of perceived effectiveness of political institutions and benefits of EU membership) and “Attachment” (participation in the election and support to the redistributive role of the EU), is confirmed in the analysis of the nine regional case studies, as they play a fundamental role in determining a region's group membership. The awareness of EU Cohesion Policy is highest in Group 2 regions Sud-Est (RO), Dolnoslaskie and Warminsko-Mazurskie (PL) and Extremadura (ES), and lowest in Group 1 “Lower EU identification” regions Essex (UK) and Norra Mellansverige (SE).

“Attachment” follows an analogue pattern. Indeed, citizens of the regions belonging to Group 2 are those who are most likely to have voted in both last elections for the European parliament, and who show a higher level of support for the EU redistributive policy.

The “Evaluation” dimension, on the other hand, is the least coherent of the three. Indeed Group 2 have the highest probability of finding citizens who consider the EU to be effective and less affected by corruption than their national institutions but, in the other groups, this dimension is more ambiguous.

The *IdentEU* results help to shed light on the patterns of EU individual and regional identification with the European project, as well as their drivers. In particular, our results are consistent with the latest tendencies emerged in the EU, i.e. the growing Euro-Skepticism that boomed with the Brexit referendum in the UK and was remarked by the results of the recent elections in France, Hungary, Italy. Indeed, it emerges that the most hostile regions to the EU project are somehow the richest ones. On the other hand, we find a high level of trust in EU institutions, in its transparency and effectiveness in Central and Eastern European countries (Hungary excluded), which instead perceive high internal level of corruption (see for instance the recent street protests in Romania and Poland).

The influencing variables that mostly affect citizens and regions' identification with the European project are those currently driving the discussion on the challenges for reforming the EU, i.e. trust in the EU institutions, the effectiveness of EU Cohesion Policy and spending, and the level of corruption. Finally, the consequences of globalization on lower skilled, lower educated and poorer workers affect citizens' identification with the European project in a negative way. Indeed, workers left worse off from the globalization show lower trust and identification with the EU institutions.

These issues are relevant at the light of three main challenges that affected the EU socio-economic development path in the last decade, i.e. the 2008 financial crisis, the globalization process, and Brexit.

First, the financial crisis led to a persisting lack of aggregate demand in most European countries, thus worsening the living conditions of the lower and middle-income classes, in particular in Southern Europe, where governments could not invest in the economic recovery as needed, being subject to tight fiscal and budgetary constraints (e.g. the fiscal compact). Then, this, in turn, might have contributed to accrue the gap in socio-economic conditions, and the trade imbalances, between the EU richer regions and the EU convergence regions.

The negative effects of the financial crisis on citizens' perception and identification with the European project (see D1.1 of the PERCEIVE project) emerged also in the focus group organized with the Local Managing Authorities within the PERCEIVE project, in particular in Emilia-Romagna, which is one of the best economic performing regions in Italy (see Brasili et al. 2017).

Second, the process of withdrawal of the United Kingdom from the European Union, known as Brexit, which started after the referendum in June 2016, increasingly led citizens (and several political parties) to question the role of the European institutions. This, in turn, contributed to increase the mistrust into the European project of the *losers* of the globalization and financial crisis, and to fuel feelings of Euro-skepticism, which eventually turn into anti-EU and anti-establishment votes in national elections in several EU member states (see France, Germany and Italy).

Third, there is growing concern that a mismanaged process of globalization contributed to decrease EU citizens' trust in the EU, as a consequence of increased inequality across households and sector, jobs' loss or delocalization, lower wages' purchasing power and booming profits on financial markets (Stiglitz, 2017; Tridico 2017). In a recent Special Eurobarometer Report, 38% of EU respondents considered globalization as a threat to employment and companies in their country, and 63% agreed that globalization increases social inequalities (Special Eurobarometer 461, 2017). In addition, the Brexit vote has been considered as a reaction to UK citizens' perception of the EU failure to protect its population from the socio-economic challenges emerged during globalization (Elliott 2016).

These insights, in particular as regards the impact of growing inequality on the identification with the EU project, are consistent with the results of D2.2, which clarified that identification with Europe increases with individuals' support to income redistribution and with trust. In addition, D2.3's results showed that countries with higher levels of the Composite Index of European Identity are those who do not need a massive support to redistribution policies and have high level of trust in their institutions.

Therefore, the *IdentEU* model not only represents an advancement on the modelling state of the art for academics, but its results also contribute to fill a major knowledge gap and to explain the drivers of the current instability and the critical phase of the EU to institutions, policy makers and practitioners at both the EU and regional level. In the ongoing debate on the need to reform EU institutions to increase its transparency and accountability to the EU citizens, our results provide a precious snapshot of EU citizens and regions' perception of the European project, and their main sources of discontent.

The results obtained through the model confirm the need to partially redirect the Cohesion Policy from the place based approach to the improvement of citizens' wellbeing, and the need to foster the integration of political interventions. This result clearly emerged from the practitioners' point of view and the communication of the Cohesion Policy carried out in the focus groups (D1.1).

With respect to further advancements of the PERCEIVE project, the results of the *IdentEU* model will be used as inputs for i) D2.5, as regards the analysis of regional disparities in citizens' identification with the European project, and their evolution over time, and ii) D6.2-d, i.e. the calibration of the System Dynamics computer model on empirical data in order to

understand the extent to which simulated behaviours are plausible in the light of available empirical evidence.

## 6. References

- Barberio, V., Mollona, E., Pareschi, L. (2017), Short contribution (report) to be used in dissemination events about the empirical relevance of a social constructivist and discursive approach to EU identity emergence and integration. Deliverable 5.1 of the PERCEIVE project. Available at [http://www.perceiveproject.eu/wp-content/uploads/2017/04/Del-5.1-Barberio\\_Mollona\\_Pareschi.pdf](http://www.perceiveproject.eu/wp-content/uploads/2017/04/Del-5.1-Barberio_Mollona_Pareschi.pdf)
- Bauhr, M. and Charron N. (2018), What do citizens think about redistribution and aid within the EU? Description and highlights of a Pan-European citizen survey on public support for Cohesion Policy, QoG Working Paper Series 2018.2 (2018): 2
- Bergbauer, S. (2018), *Explaining European Identity Formation*, Springer: Netherlands.
- Brasili, C., Aiello V., P. Calia, P. M. Reverberi (2017), D1.1 “Report on regional case-studies”, PERCEIVE project, GA no. 693529, [www.perceiveproject.eu](http://www.perceiveproject.eu)
- Braun, D. and M. Tausendpfund (2014), The impact of the Euro crisis on citizens’ support for the European Union. *Journal of European Integration*, 36, pp. 231–245.
- Bruter, M. (2009), Time bomb – news, media, and European identity. *Comparative Political Studies*, 42 (12).
- Bruter, M. (2008), Legitimacy, Euroscepticism & Identity in the European Union – Problems of Measurement, Modelling & Paradoxical Patterns of Influence’, *Journal of Contemporary European Research*, Vol. 4, No. 4, pp. 273-285.
- Bruter, M. (2005), *Citizens of Europe? The Emergence of a Mass European Identity*. Basingstoke: Palgrave.
- Bruter, M. (2004), On what citizens mean by feeling European. *Journal of Ethnic and Migration Studies*, 30 (1), pp. 21-41.
- 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), pp. 1148–79.
- Boijmans, P. (2014), Administrative Capacity Building linked to the management of ESI Funds Priorities for 2014-2016, Committee of the Regions, 19 June 2014
- Charron, N. (2017), Report including the database development and implementation and how it supports all the project WPs, Deliverable 2.1, PERCEIVE Project (GA nr. 693529).
- Charron N., Dijkstra L., Lapuente V. (2014), Regional Governance Matters: Quality of Government within European Union Member States, *Regional Studies*, 48:1, 68-90
- Checkel, J. T., Katzenstein, P. J. (2009), The politicization of European identities. In J. Checkel, & P. Katzenstein (Eds), *European identity* (pp. 1-28). Cambridge: Cambridge University Press.
- Citrin, J., Sides, J. (2004), More than nationals: How identity choice matters in the new Europe. In R. Herrmann, T. Risse, M. Brewer (Eds), *Transnational Identities: Becoming European in the EU* (pp. 161-185). Lanham: Rowman & Littlefield.
- David, O., and Bar-Tal, D. (2009), A sociopsychological conception of collective identity: The case of national identity as an example. *Personality and Social Psychology Review*, 13(4), pp. 354-379.
- Easton, D. (1965), *A systems analysis of political life*. New York: Wiley.
- Elliott, L. (2016), Brexit is a rejection of globalization. *Financial Times*, 26 June 2016.
- Eurobarometer (2017), *Designing Europe’s future Report*. Special Eurobarometer 461, April 2017.
- Fligstein, N., Polyakova, A. and Sandholtz, W. (2012), European integration, nationalism and European identity. *Journal of Common Market Studies*, 50(S1), pp. 106-122.
- Fligstein, N. (2008), *Euroclash: The EU, European identity, and the future of Europe*. Oxford: Oxford University Press.

- Gabel, M. and Palmer, H. (1995), Understanding Variation in Public Support of European Integration. *European Journal of Political Research*, 27, pp. 3–19.
- Grundy, S. and Jamieson, L. (2007), European Identities: from absent-minded citizens to passionate Europeans. *Sociology* 41 (4), pp. 663–680
- Haas, E. (1968), *The Uniting of Europe*, Notre Dame, IN: Notre Dame University Press.
- Haas, E. (1961), International Integration: The European and the Universal Process. *International Organization*, 15, pp. 366–92.
- Habermas, J. (2012), The crisis of the European Union: A response. *Polity*.
- Habermas, J. (2011), *Zur Verfassung Europas. Ein Essay*. Berlin: Suhrkamp.
- Habermas, J. (2001), Why Europe Needs a Constitution. *New Left Review*, II 5–26.
- Hagenaars, J. A. and McCutcheon, A. L. (2002), *Applied Latent Class Analysis*, Cambridge University Press, Cambridge.
- Harteveld, E., T. van der Meer and C.E. de Vries (2013), In Europe we trust? Exploring three logics of trust in the European Union. *European Union Politics*, 14, pp. 542–565.
- Herrmann, R., Brewer, M. (2004), Why study identity and institutions. In R. Herrmann, T. Risse & M. Brewer (Eds), *Transnational Identities: Becoming European in the EU* (pp. 1–10). Lanham: Rowman & Littlefield.
- Herrmann, R., Risse, T. and Brewer, M. (eds) (2004), *Transnational Identities*. Boulder: Rowman & Littlefield.
- Holzhacker, R. (2007), Democratic legitimacy and the European Union. *European Integration*, 29(3), pp.257–269.
- Hooghe, M. and Verhaegen, S. (2017), The effect of political trust and trust in European citizens on European identity. *European Political Science Review*, 9(2), pp.161–181.
- Lopez-Bazo, E. and Royuela, V. (2017), Mapping the determinants of EU citizen's perception and identification. Deliverable 2.2 PERCEIVE project. Available at [http://www.perceiveproject.eu/wp-content/uploads/2017/09/PERCEIVE-Project-Deliverable-2.2\\_Mapping-the-determinants-of-EU-citizen%E2%80%99s-perceptionand-identification.pdf](http://www.perceiveproject.eu/wp-content/uploads/2017/09/PERCEIVE-Project-Deliverable-2.2_Mapping-the-determinants-of-EU-citizen%E2%80%99s-perceptionand-identification.pdf)
- Luhman, S. (2017), The European Identity Survey – a Bridge between political science and psychology. *Journal of Contemporary European Research*, 13(3), pp. 1261–1282.
- Marcussen, M., Risse, T., Engelmann-Martin, D., Knopf, H., and Roscher, K. (1999), Constructing Europe? The evolution of French, British and German nation state identities. *Journal of European Public Policy*, 6(4), 614–633.
- Meinhof, U., Galasinski, D. (2005), *The language of belonging*. Basingstoke, UK: Macmillan.
- Meyer, R., Hammerschmid, G. (2006), Changing institutional logics and executive identities: A managerial challenge to public administration in Austria. *American Behavioral Scientist*, 49(7), 1000–1014.
- Mendez, C. and Batchler, J. (2017), European Identity and Citizen Attitudes to Cohesion Policy: What do we know? COHESIFY RESEARCH PAPER 1.
- Osterloh, S. (2011), Can regional transfers buy public support? Evidence from EU structural policy. Working paper, available at SSRN <https://goo.gl/T2EbEB>
- 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.
- Pirani, E. (2013), Evaluating contemporary social exclusion in Europe: a hierarchical latent class approach, *Quality and Quantity*, 47, 923–941.
- Reeskens, T. and Hooghe, M. (2010), Beyond the Civic–Ethnic Dichotomy. *Nations and Nationalism*, 16, pp. 579–97.
- Risse, T. (2014), No demos? Identities and public spheres in the Euro Crisis. *Journal of Common Market Studies*, 52(6), pp. 1207–1215.

- Risse, T. (2010), *A Community of Europeans? Transnational Identities and Public Spheres*. Ithaca: Cornell University Press.
- Rodríguez-Pose, A., Tselios, V. (2011), Mapping the European regional educational distribution. *European Urban and Regional Studies*, 18(4), pp. 358-374.
- Rodríguez-Pose, A., Tselios, V. (2010), Inequalities in income and education and regional economic growth in Western Europe. *The Annals of Regional Science*, 44(2), pp. 349–375
- Rodríguez-Pose, A., Tselios, V. (2009), Education and income inequality in the regions of the European Union. *Journal of Regional Science*, 49(3), pp. 411–437.
- Royuela, V. (2018), Report on the construction of the CIEI indicator. Qualitative report with main findings from the survey and discussion of comparative results from the application of the indicator of perception in the different case study regions, Deliverable 2.3, PERCEIVE Project (GA nr. 693529)
- Skrondal, A. and Rabe-Hesketh, S. (2004), *Generalized latent variable modeling: multilevel, longitudinal and structural equation models*, Chapman and Hall/CRC, Boca Raton.
- Stiglitz, J.E. (2017), *Globalization and Its Discontents Revisited: Anti-Globalization in the Era of Trump*. W. W. Norton & Company.
- Tajfel, H. and Turner, J.C. (1986), *The Social Identity Theory of Inter-group Behaviour*. In Worchel, S. and Austin, L.W. (eds) *Psychology of Intergroup Relations*, Chicago, IL, Nelson-Hall.
- Tajfel, H. (1981), *Human Groups and Social Categories*, Cambridge, Cambridge University Press.
- Tosun, J. (2013), *Absorption of Regional Funds: A comparative Analysis*, JCMS, 2013
- Tridico, P. (2017), *Inequality in financial capitalism*. Taylor & Francis.
- Tucker, P. (2018), *Unelected Power: The Quest for Legitimacy in Central Banking and the Regulatory State*. Princeton University Press.
- Verhaegen, S., Hooghe, M. and Quintelier, E. (2014), European Identity and Support for European Integration: A Matter of Perceived Economic Benefits?. *Kyklos*, 64(2), pp. 295–314.
- Vermunt, J. K. (2003), Multilevel Latent Class Models, *Sociological Methodology*, 33, 213–239.
- Vermunt J.K. and Magidson J. (2016), *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*, Belmont, MA: Statistical Innovations Inc.



## 7. Appendix

**Table A.1.** Complementarities between D2.4 and the other PERCEIVE deliverables

Deliverable	WP	Title	Focus	Complementarity (input/output)
D1.1 (UNIBO)	WP1	Report on regional case-studies.	Focus groups with Local Managing Authorities.	Input: insights for selection of model's clusters and contextualization of <i>IdentEU</i> model's results.
D1.2 (UGOT)	WP1	Dataset built from the survey at citizen level for the case-studies regions and report with qualitative results.	PERCEIVE survey and PERCEIVE EU Regional Dataset.	Input: data from the PERCEIVE Survey at the individual level, and at the regional level from the PERCEIVE's EU Regional Dataset (including indicators of the Cohesion Policy in each region) are used to build the <i>IdentEU</i> probabilistic model.
D2.2 (UBO)	WP2	Mapping the determinants of EU citizens' perception and identification.	Analysis of the determinants of citizens' identification with the EU.	Input: use of the results on citizens' identification with the EU project, awareness and perception of the Cohesion policies to select the models' variables.
D2.3 (UBO)	WP2	Report on the construction of the CIEI indicator.	Development of the Composite Index of European Identity.	Input: use of the Composite Index of European Identity in the model development.
D5.1 (WU, UNIBO)	WP5	Short contribution to be used in dissemination events about the empirical relevance of a social constructivist and discursive approach to EU identity emergence and integration.	Framework of analysis and the literature review on the definition and the determinants of European identity.	Input: determinants of European identity and identification with the EU to inform the selection of the model's variables.
D2.5 (UBO, UNIBO)	WP2	Analysis of regional disparities in EU citizens' identification and its evolution over time.	Identification of regional patterns of disparities in citizens' identification with the EU project with spatial analyses.	Output: results of the <i>IdentEU</i> model as regards the drivers of citizens and regions' identification with the EU.
D2.6 (UNIBO, IEA, IAFE-NRI, UB, PBS, WU, BAM!)	WP6	Quantitative model building and validation	Calibration of the System Dynamics computer model on empirical data in order to understand the extent to which simulated behaviours are plausible in the light of available empirical evidence.	Output: results of the <i>IdentEU</i> model in terms of variables influencing citizens and regions' identification with the EU.

D2.4 uses inputs from D1.1, D1.2, D2.2, D2.3, D5.1 to identify the determinants of identity and identification with the EU to be quantified and assessed by *IdentEU*. In contrast, *IdentEU* results (in particular on the drivers of citizens and regions' identification with the EU) feed into D2.5 and D2.6d.

**Table A.2.** Definition of indicators, individual covariates and contextual variables

Variable	Description	Values
<b>Indicators (manifest variables)</b>		
Awareness of EU policies	Q1. Have you ever heard about the following EU policies? (EU Cohesion Policy; EU Regional Policy; EU Structural Funds; any EU funded project in your region or area)	0 None of these 1 Only local project 2 At least one among EU CP, EU RP, EU SF
Benefits from EU policies	Q3. Have you ever benefited in your daily life from any project funded by the EU?	1 Yes 2 No 99 Don't Know/RF
Identification with Europe (Q9_3)	Q9. On a 0-10 scale, with '0' being 'I don't identify at all' and '10' being "I identify very strongly", how strongly you identify yourself with the following: Q9_1: Your region; Q9_2: Your country; Q9_3: Europe	1 Not much (0-3) 2 Somewhat strongly (4-6) 3 Strongly (7-10)
Identification with Europe vs Country	Comparing Q9_3 to Q9_1	1 Less 2 Equal 3 More
Identification with Europe vs Region	Comparing Q9_3 to Q9_2	1 Less 2 Equal 3 More
Effectiveness of EU (Q5_1)	Q5. How effective do you think the following institutions will be at dealing with the biggest problem in your region? Q5_1: The EU; Q5_2: National governing institutions; Q5_3: Regional/local Institutions	1 Not very effective 2 Somewhat effective 3 Very effective
Effectiveness of EU vs National governing institutions	Comparing Q5_1 to Q5_2	1 Less 2 Equal 3 More
Effectiveness of EU vs regional/local governing institutions	Comparing Q5_1 to Q5_3	1 Less than EU 2 Equal to EU 3 More than EU
Evaluation of EU membership	Q8. In general, do you think that (YOUR COUNTRY'S) EU membership is a good thing, a bad thing, neither good nor bad?	1 Good 0 Bad / Neither good or bad/ Not sure
Corruption in EU (Q16_1)	Q16. On a 0-10 scale, with '0' being that 'there is no corruption' and '10' being that corruption is widespread, how would you rate: Q16_1: The European union; Q16_2: The national government; Q16_3: The region/local government?	1 Low (0-3) 2 Medium (4-6) 3 High (7-10)
Corruption in EU vs National government	Comparing Q16_1 to Q16_2	1 Less 2 Equal 3 More
Corruption in EU vs regional/local government	Comparing Q16_1 to Q16_3	1 Less 2 Equal 3 More

**Table A.2. (Continue)**

Variable	Description	Values
<b>Indicators</b>		
Vote in the EU elections	Q7. Have you voted in either of the last two EU parliamentary elections?	0 Neither 1 Once 3 Both 999 Don't know/RF
Support	Q20: In your opinion, the EU should continue this policy, where wealthier countries contribute more, and poorer EU regions receive more funding?	1 – Agree; 2 – Disagree; 3 – D/K
<b>Individual Covariates</b>		
Gender	Gender	1 Male 2 Female
Age	Age in years	Continuous
Education	Level of education	1 - Up to first level secondary 2 - High school 3 - Degree and PhD
Level of income	Net income per month (after taxes)	1 - Low 2 - Medium 3 - High
Occupation	Occupational status	1 - Employee 2 – unemployed 3 - Housewife, pensioner, retired, Other 4 Students, trainee
<b>Context covariates (at regional level)</b>		
GDP_14	GDP per inhabitant in Euro at 2014 as percentage of EU average	Numerical value
Eqi_100	European Index of Institutional Quality at 2013 (Normalized)	Numerical value (0-100)
Abs_13	Absorption rate of SF expenditures: ratio of SF expenditures up to 2013 to the SF allocation in the 2007-2013 period	Numerical value (0-1)
SFtot_pc	Total expenditures financed by Structural Fund in years 2007-13 divided by the average population in a region in the period 2007-13	Numerical value (euros)
Unold14	Unemployment rate: population 20-64 years	Numerical value

### A.3. Bic statistics and number of parameter for models with T individual clusters and M regions' groups

	Regions' Group									
	1		2		3		4		5	
	<i>BIC</i>	<i>Npar</i>	<i>BIC</i>	<i>Npar</i>	<i>BIC</i>	<i>Npar</i>	<i>BIC</i>	<i>Npar</i>	<i>BIC</i>	<i>Npar</i>
<b>Individual Clusters</b>										
1	298870	22	298108	28	298010	34	297992	40	297997	46
2	284111	49	283256	56	283059	63	282986	70	282965	77
3	279016	76	277771	84	277307	92	277058	100	277013	108
4	276010	103	274491	112	274243	121	273666	130	273458	139
5	273440	130	271812	140	271693	150	271284	160	271367	170
6	271789	157	269925	168	269578	179	269301	190	268971	201
7	270153	184	268337	196	268481	208	268228	220	267702	232

**Table A.4.** Prior cluster membership probabilities  $P(W = m | \mathbf{Z})$ , and posterior cluster membership probabilities  $P(W = m | \mathbf{Y}, \mathbf{Z})$  with the respective modal assignment for the regions.

Region	Posterior Probability					Model Probability				
	Modal	1	2	3	4	Modal	1	2	3	4
DE1 - Baden-Württemberg	3	0.00	0.00	1.00	0.00	3	0.28	0.00	0.72	0.00
DE2 - Bayern	3	0.00	0.00	1.00	0.00	3	0.20	0.00	0.79	0.00
DE3 - Berlin	3	0.00	0.00	1.00	0.00	1	0.80	0.04	0.04	0.12
DE4 - Brandenburg	3	0.00	0.00	1.00	0.00	3	0.29	0.07	0.62	0.03
DE5 - Bremen	3	0.10	0.00	0.90	0.00	3	0.41	0.00	0.55	0.04
DE6 - Hamburg	3	0.00	0.00	1.00	0.00	3	0.20	0.00	0.79	0.02
DE7 - Hessen	3	0.00	0.00	1.00	0.00	3	0.38	0.00	0.62	0.00
DE8 - Mecklenburg-Vorpommern	1	0.74	0.00	0.26	0.00	2	0.34	0.39	0.19	0.08
DE9 - Niedersachsen	3	0.00	0.00	1.00	0.00	1	0.53	0.00	0.47	0.01
DEA - Nordrhein-Westfalen	3	0.00	0.00	1.00	0.00	1	0.70	0.00	0.28	0.01
DEB - Rheinland-Pfalz	3	0.00	0.00	1.00	0.00	3	0.49	0.00	0.51	0.00
DEC - Saarland	3	0.05	0.00	0.95	0.00	1	0.60	0.00	0.39	0.01
DED - Sachsen	3	0.00	0.00	1.00	0.00	3	0.22	0.12	0.38	0.28
DEE - Sachsen-Anhalt	3	0.00	0.00	1.00	0.00	2	0.22	0.30	0.24	0.24
DEF - Schleswig-Holstein	3	0.01	0.00	0.99	0.00	1	0.56	0.00	0.44	0.00
DEG - Thüringen	3	0.00	0.00	1.00	0.00	3	0.20	0.04	0.74	0.02
SE11 - Stockholm	1	0.99	0.00	0.01	0.00	1	0.51	0.00	0.49	0.01
SE21 - Småland med öarna	1	0.99	0.00	0.01	0.00	1	0.81	0.00	0.17	0.02
SE31 - Norra Mellansverige	1	1.00	0.00	0.00	0.00	1	0.61	0.00	0.37	0.02
AT11 - Burgenland	3	0.01	0.00	0.99	0.00	3	0.31	0.01	0.68	0.00
AT12 - Niederösterreich	3	0.00	0.00	1.00	0.00	1	0.66	0.00	0.33	0.01
AT13 - Wien	1	0.75	0.00	0.25	0.00	1	0.55	0.00	0.02	0.42
AT21 - Kärnten	1	0.89	0.00	0.00	0.11	1	0.74	0.01	0.22	0.03
AT22 - Steiermark	3	0.02	0.00	0.98	0.00	1	0.56	0.00	0.43	0.00
AT31 - Oberösterreich	3	0.00	0.00	1.00	0.00	3	0.41	0.00	0.58	0.01
AT32 - Salzburg	3	0.01	0.00	0.99	0.00	3	0.22	0.00	0.78	0.00
AT33 - Tirol	3	0.00	0.00	1.00	0.00	3	0.29	0.00	0.70	0.01
AT34 - Vorarlberg	3	0.11	0.00	0.89	0.00	3	0.22	0.00	0.78	0.00
BG31 - Severozapaden	2	0.00	1.00	0.00	0.00	2	0.04	0.96	0.00	0.01
BG32 - Severen centralen	2	0.00	1.00	0.00	0.00	2	0.06	0.93	0.00	0.01
BG33 - Severoiztochen	1	1.00	0.00	0.00	0.00	2	0.09	0.91	0.00	0.01
BG34 - Ugoiztochen	1	0.99	0.00	0.00	0.01	2	0.14	0.86	0.00	0.00
BG41 - Ugozapaden	2	0.00	1.00	0.00	0.00	1	0.52	0.47	0.00	0.01
BG42 - Uzhen centralen	2	0.00	1.00	0.00	0.00	2	0.10	0.89	0.00	0.00
EE00 - Estonia	1	1.00	0.00	0.00	0.00	2	0.01	0.50	0.35	0.14
ES11 - Galicia	2	0.00	1.00	0.00	0.00	2	0.00	0.89	0.00	0.11
ES12 - Principado de Asturias	2	0.00	1.00	0.00	0.00	2	0.01	0.98	0.00	0.01
ES13 - Cantabria	2	0.01	0.96	0.00	0.03	2	0.04	0.83	0.00	0.13
ES21 - Pais Vasco	2	0.00	1.00	0.00	0.00	1	0.54	0.37	0.00	0.09
ES22 - Comunidad Foral de Navarra	2	0.00	1.00	0.00	0.00	1	0.39	0.36	0.00	0.25

ES23 - La Rioja	2	0.28	0.72	0.00	0.00	2	0.44	0.55	0.00	0.01
ES24 - Aragón	2	0.00	1.00	0.00	0.00	2	0.05	0.66	0.00	0.29
ES30 - Comunidad de Madrid	2	0.00	1.00	0.00	0.00	1	0.69	0.30	0.00	0.01
ES41 - Castilla y León	2	0.00	1.00	0.00	0.00	2	0.03	0.92	0.00	0.05
ES42 - Castilla-La Mancha	2	0.00	1.00	0.00	0.00	2	0.00	1.00	0.00	0.00
ES43 - Extremadura	2	0.00	1.00	0.00	0.00	2	0.00	0.96	0.00	0.04
ES51 - Cataluña	2	0.00	1.00	0.00	0.00	2	0.13	0.74	0.00	0.13
ES52 - Comunidad Valenciana	2	0.00	1.00	0.00	0.00	2	0.03	0.97	0.00	0.00
ES53 - Illes Balears	2	0.00	1.00	0.00	0.00	2	0.12	0.82	0.00	0.06
ES61 - Andalucía	2	0.00	1.00	0.00	0.00	2	0.00	0.68	0.00	0.32
ES62 - Región de Murcia	2	0.00	1.00	0.00	0.00	2	0.00	1.00	0.00	0.00
ES70 - Canarias (ES)	2	0.00	1.00	0.00	0.00	2	0.00	0.93	0.00	0.07
FR10 - Ile-de-France	1	1.00	0.00	0.00	0.00	1	0.68	0.00	0.08	0.24
FR21 - Champagne-Ardenne	1	0.92	0.02	0.06	0.00	1	0.85	0.10	0.01	0.04
FR22 - Picardie	1	0.98	0.00	0.02	0.00	1	0.59	0.25	0.00	0.15
FR23 - Haute-Normandie	1	0.76	0.00	0.00	0.24	4	0.33	0.13	0.00	0.54
FR24 - Centre	1	1.00	0.00	0.00	0.00	1	0.80	0.09	0.01	0.10
FR25 - Basse-Normandie	1	0.99	0.00	0.00	0.00	1	0.82	0.07	0.03	0.08
FR26 - Bourgogne	1	1.00	0.00	0.00	0.00	1	0.74	0.13	0.01	0.12
FR30 - Nord - Pas-de-Calais	1	1.00	0.00	0.00	0.00	1	0.55	0.33	0.00	0.12
FR41 - Lorraine	1	0.91	0.00	0.00	0.09	1	0.52	0.28	0.00	0.19
FR42 - Alsace	1	0.90	0.00	0.00	0.10	1	0.72	0.06	0.01	0.21
FR43 - Franche-Comte	1	1.00	0.00	0.00	0.00	1	0.76	0.16	0.01	0.07
FR51 - Pays de la Loire	1	1.00	0.00	0.00	0.00	1	0.85	0.04	0.03	0.07
FR52 - Bretagne	1	1.00	0.00	0.00	0.00	1	0.89	0.03	0.07	0.01
FR53 - Poitou-Charentes	1	1.00	0.00	0.00	0.00	1	0.71	0.15	0.01	0.14
FR61 - Aquitaine	1	1.00	0.00	0.00	0.00	1	0.86	0.06	0.03	0.06
FR62 - Midi-Pyrenees	1	1.00	0.00	0.00	0.00	1	0.78	0.06	0.02	0.14
FR63 - Limousin	1	0.99	0.00	0.00	0.00	1	0.83	0.11	0.02	0.04
FR71 - Rhone-Alpes	1	1.00	0.00	0.00	0.00	1	0.92	0.01	0.05	0.01
FR72 - Auvergne	1	1.00	0.00	0.00	0.00	1	0.84	0.05	0.07	0.04
FR81 - Languedoc-Roussillon	1	1.00	0.00	0.00	0.00	1	0.69	0.24	0.00	0.07
FR82 - Provence-Alpes-Cote d'Azur	1	1.00	0.00	0.00	0.00	1	0.63	0.06	0.01	0.31
FR83 - Corse	4	0.02	0.00	0.00	0.98	4	0.25	0.12	0.01	0.62
HU10-Central Hungary	4	0.00	0.00	0.00	1.00	1	0.56	0.23	0.19	0.02
HU21-Central Transdanubia	4	0.00	0.00	0.00	1.00	4	0.00	0.26	0.13	0.61
HU22-Western Transdanubia	4	0.00	0.00	0.00	1.00	4	0.00	0.12	0.16	0.73
HU23-Southern Transdanubia	4	0.00	0.00	0.00	1.00	4	0.00	0.35	0.01	0.64
HU31-Northern Hungary	4	0.00	0.00	0.00	1.00	4	0.00	0.25	0.00	0.75
HU32-Northern Great Plain	4	0.00	0.00	0.00	1.00	4	0.00	0.26	0.00	0.74
HU33-Southern Great Plain	4	0.00	0.00	0.00	1.00	4	0.00	0.23	0.00	0.77
LV00-Latvia	4	0.00	0.00	0.00	1.00	2	0.00	0.81	0.00	0.19
NL11-Groningen	1	1.00	0.00	0.00	0.00	1	0.63	0.00	0.36	0.00
NL12-Friesland	1	1.00	0.00	0.00	0.00	1	0.87	0.01	0.13	0.00



NL13-Drenthe	1	0.75	0.00	0.24	0.00	1	0.80	0.00	0.20	0.00
NL21-Overijssel	1	0.99	0.00	0.01	0.00	1	0.79	0.00	0.21	0.00
NL22-Gelderland	1	1.00	0.00	0.00	0.00	1	0.77	0.00	0.22	0.00
NL23-Flevoland	1	0.99	0.00	0.00	0.01	1	0.89	0.05	0.01	0.05
NL31-Utrecht	1	1.00	0.00	0.00	0.00	1	0.58	0.00	0.38	0.04
NL32-Noord-Holland	1	1.00	0.00	0.00	0.00	1	0.59	0.00	0.31	0.10
NL33-Zuid-Holland	1	1.00	0.00	0.00	0.00	1	0.71	0.01	0.06	0.23
NL34-Zeeland	1	0.96	0.00	0.04	0.00	1	0.51	0.00	0.49	0.00
NL41-Noord-Brabant	1	1.00	0.00	0.00	0.00	1	0.56	0.00	0.44	0.00
NL42-Limburg	1	1.00	0.00	0.00	0.00	1	0.83	0.00	0.16	0.00
PL11 – Lodzkie	2	0.00	1.00	0.00	0.00	2	0.01	0.83	0.01	0.15
PL12 - Mazowieckie	2	0.00	1.00	0.00	0.00	2	0.04	0.43	0.14	0.39
PL21 - Malopolskie	2	0.00	1.00	0.00	0.00	2	0.01	0.85	0.01	0.13
PL22 – Slaskie	2	0.00	1.00	0.00	0.00	2	0.02	0.85	0.02	0.11
PL31 - Lubelskie	2	0.00	1.00	0.00	0.00	2	0.00	0.91	0.00	0.09
PL32 - Podkarpackie	2	0.00	1.00	0.00	0.00	2	0.00	0.93	0.00	0.07
PL33 - Swietokrzyskie	2	0.00	0.99	0.00	0.01	2	0.00	0.95	0.00	0.04
PL34 - Podlaskie	2	0.00	1.00	0.00	0.00	2	0.00	0.83	0.00	0.16
PL41 - Wielkopolskie	2	0.00	1.00	0.00	0.00	2	0.03	0.78	0.06	0.13
PL42 - Zachodniopomorskie	2	0.00	1.00	0.00	0.00	2	0.01	0.87	0.02	0.10
PL43 - Lubuskie	2	0.00	1.00	0.00	0.00	2	0.02	0.89	0.03	0.06
PL51 - Dolnoslaskie	2	0.00	1.00	0.00	0.00	2	0.01	0.83	0.01	0.14
PL52 - Opolskie	2	0.00	1.00	0.00	0.00	2	0.01	0.83	0.02	0.14
PL61 - Kujawsko-Pomorskie	2	0.00	1.00	0.00	0.00	2	0.00	0.89	0.00	0.11
PL62 - Warminsko-Mazurskie	2	0.00	1.00	0.00	0.00	2	0.00	0.91	0.00	0.09
PL63 - Pomorskie	2	0.00	1.00	0.00	0.00	2	0.01	0.85	0.01	0.12
RO11 -Nord-Vest	2	0.00	1.00	0.00	0.00	2	0.16	0.63	0.12	0.09
RO12 –Centru	2	0.00	1.00	0.00	0.00	2	0.02	0.79	0.00	0.19
RO21 -Nord-Est	2	0.00	1.00	0.00	0.00	2	0.11	0.76	0.05	0.08
RO22 -Sud-Est	2	0.00	1.00	0.00	0.00	2	0.01	0.84	0.00	0.15
RO31 -Sud-Muntenia	2	0.00	1.00	0.00	0.00	2	0.02	0.85	0.00	0.12
RO32 -Bucure?ti-Ilfov	2	0.00	1.00	0.00	0.00	4	0.04	0.16	0.01	0.79
RO41 -Sud-Vest Oltenia	2	0.00	1.00	0.00	0.00	2	0.05	0.85	0.01	0.09
RO42 –Vest	2	0.00	1.00	0.00	0.00	2	0.15	0.65	0.07	0.13
SK01 -Bratislava Region	2	0.00	1.00	0.00	0.00	4	0.20	0.01	0.29	0.49
SK02 -Western Slovakia	2	0.00	1.00	0.00	0.00	4	0.00	0.44	0.00	0.56
SK03 -Central Slovakia	2	0.00	1.00	0.00	0.00	4	0.00	0.50	0.00	0.50
SK04 -Eastern Slovakia	2	0.00	1.00	0.00	0.00	2	0.00	0.64	0.00	0.36
UKC - North East, England	1	1.00	0.00	0.00	0.00	1	0.88	0.05	0.04	0.02
UKD - North West, England	1	1.00	0.00	0.00	0.00	1	0.72	0.01	0.27	0.00
UKE - Yorkshire and the Humber	1	1.00	0.00	0.00	0.00	1	0.85	0.02	0.11	0.02
UKF - East Midlands, England	1	1.00	0.00	0.00	0.00	1	0.78	0.01	0.20	0.01
UKG - West Midlands, England	1	1.00	0.00	0.00	0.00	1	0.86	0.02	0.09	0.03
UKH - East of England	1	1.00	0.00	0.00	0.00	1	0.68	0.00	0.32	0.01
UKI - London, England	1	1.00	0.00	0.00	0.00	3	0.34	0.00	0.52	0.14

UKJ - South East, England	1	1.00	0.00	0.00	0.00	3	0.49	0.00	0.50	0.01
UKK - South West, England	1	1.00	0.00	0.00	0.00	1	0.50	0.00	0.49	0.00
UKL - Wales	1	0.58	0.00	0.00	0.42	1	0.39	0.09	0.31	0.20
UKM - Scotland	1	1.00	0.00	0.00	0.00	1	0.66	0.00	0.33	0.01
UKN - Northern Ireland	1	0.99	0.00	0.00	0.01	1	0.75	0.02	0.21	0.01
ITC1 - Piemonte	4	0.00	0.00	0.00	1.00	1	0.79	0.13	0.01	0.08
ITC3 - Liguria	1	0.97	0.00	0.00	0.03	1	0.85	0.09	0.01	0.05
ITC4 - Lombardia	4	0.00	0.00	0.00	1.00	1	0.91	0.01	0.06	0.02
ITF1 - Abruzzo	1	0.95	0.00	0.00	0.05	1	0.70	0.27	0.00	0.02
ITF3 - Campania	4	0.00	0.00	0.00	1.00	4	0.00	0.20	0.00	0.80
ITF4 - Puglia	4	0.00	0.00	0.00	1.00	4	0.00	0.44	0.00	0.56
ITF6 - Calabria	4	0.00	0.00	0.00	1.00	4	0.00	0.37	0.00	0.63
ITG1 - Sicilia	4	0.00	0.00	0.00	1.00	4	0.00	0.49	0.00	0.51
ITG2 - Sardegna	4	0.01	0.00	0.00	0.99	2	0.01	0.66	0.00	0.33
ITH1 - Trentino-Alto Adige/Südtirol	1	0.92	0.00	0.02	0.06	3	0.40	0.00	0.58	0.01
ITH3 - Veneto	4	0.33	0.00	0.00	0.67	1	0.87	0.02	0.06	0.05
ITH4 - Friuli-Venezia Giulia	1	0.99	0.00	0.00	0.01	1	0.92	0.02	0.06	0.01
ITH5 - Emilia-Romagna	4	0.00	0.00	0.00	1.00	1	0.71	0.03	0.02	0.24
ITI1 - Toscana	4	0.00	0.00	0.00	1.00	1	0.66	0.09	0.01	0.24
ITI2 - Umbria	1	0.88	0.00	0.00	0.12	1	0.71	0.23	0.00	0.05
ITI3 - Marche	1	0.64	0.00	0.00	0.36	1	0.85	0.10	0.01	0.04
ITI4 - Lazio	4	0.00	0.00	0.00	1.00	1	0.51	0.10	0.00	0.39