XXXVIII CONFERENZA ITALIANA DI SCIENZE REGIONALI

DOES EARTHQUAKE HAZARD COUPLE WITH SOCIO-ECONOMIC VULNERABILITY? RESULTS FROM A MUNICIPALITY-LEVEL ANALYSIS IN ITALY

Margherita Russo[1], Francesco Pagliacci[2], Riccardo Righi[3]

ABSTRACT

A series of massive earthquakes struck Central Italy in 2016, following the 2009 earthquake in L'Aquila and the one in 2012 in Emilia-Romagna. Because of their large impact in terms of human losses and socio-economic uncertainty, these events spurred political decision-makers to advocate an ambitious long-term re-structuring intervention, the Piano Casa Italia. This represents a political, economic and social challenge, involving the multifaceted relationship between human settlements, human activities and territory. Thus, detecting heterogeneity of local areas should help policy-makers to fix the place-based priorities and to define some general procedures of interventions. This paper takes an analytical perspective to support such informed policy measures. This work considers socio-economic, demographic and geographic variables at a municipality level and, by applying a Principal Component Analysis, it singles out a fewer number of the most important components characterizing them. Grounded on those components, hierarchical cluster analysis is performed and cluster composition is analyzed with regard to the four seismic zones that the Italian Civil Protection defines. Data suggest that, in Italy, seismic hazard couples with socio-economic vulnerability. Indeed, the territorial allocation of the clusters that show the greatest vulnerability overlaps with the most seismic zones. This result supports a policy mix approach in reducing vulnerability and enhancing living and working conditions in those areas.

Acknowledgments
This paper has been realized in the research project Energie Sisma Emilia of the University of Modena and Reggio Emilia (Italy) (www.energie.unimore.it). The project is co-funded by Fondazione Cassa di Risparmio di Modena, Azienda Unità Sanitaria Locale di Modena and Università di Modena e Reggio Emilia.
The views expressed are those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

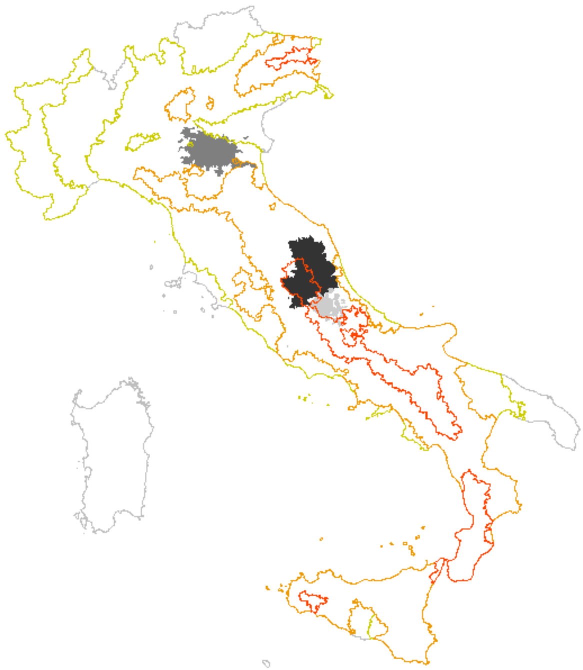
[1] DEMB, Università di Modena e Reggio Emilia, viale Berengario 51, 41121, Modena, and CAPP-DEMB UniMORE, e-mail: margherita.russo@unimore.it.
[2] DEMB, Università di Modena e Reggio Emilia, viale Berengario 51, 41121, Modena, and CAPP-DEMB UniMORE, e-mail: francesco.pagliacci@unimore.it (corresponding author).
[3] European Commission-JRC- Dir B, Seville, Spain, and CAPP-DEMB UniMORE, e-mail riccardo.righi@unimore.it

1. Introduction

Italy has a long history of earthquakes (Chubb, 2002; Guidoboni, 2012; 2014), as it "sits on two fault lines, making it one of the most seismically active countries in Europe" (UNISDR, 2015). As far as the last decade is concerned, three major earthquakes have struck the country: L'Aquila (2009), Emilia (2012), and Central Italy (2016). Each of them had large impact of human losses and economic and social uncertainty. In particular, both the 2009 L'Aquila earthquake (6.3 M_w) and the 2016 Norcia-Amatrice long series of earthquakes (6.6 M_w at a maximum) were extremely powerful, with a death toll of 309 and 299 people, respectively. They also produced massive destructions across Central Italy. The 2012 earthquake in Emilia was not as devastating as the aforementioned ones (it was characterised by two seisms, with a maximum magnitude of 6.1 Mw and a death toll hitting 27), but it caused severe damages, as well. Indeed, the affected area was amongst the country's wealthiest and most productive ones, with the presence of important industrial and agricultural districts, significantly contributing to exports (Barone et al., 2013). Comprehensively, the three aforementioned events affected hundreds of municipalities and more than 15 thousand square kilometres in total (almost 5% of total national area), with a population of 2.4 million inhabitants and 0.8 million employees, respectively 5% and 3% at country level (Figure 1).

Figure 1 – Latest earthquakes in Italy

Seismic zones	Latest earthquakes	No. Municipalities	Population (million)	Employment (million)
1	Central Italy (2016)	131	0.57	0.15
2	Emilia (2012)	109	1.71	0.59
3	L'Aquila (2009)	49	0.13	0.03
4		289	2.41	0.77



Source: authors' elaboration on Civil Protection (2015) data

Besides these specific events, Italy as a whole shows a high seismic risk. In order to reduce the impact of earthquake hazard, the Italian Civil Protection (2015) has provided a territorial classification for the whole

country, which is based on past earthquakes' intensity and frequency[4]. Such a classification shows that 36% of Italian municipalities are classified with the highest seismic risk (zones #1 and #2), (2015). They are about three thousand municipalities, accounting for 39% of the Italian population and 31% of total employment[5]. These figures provide an idea of the social and economic burden of seismic hazard for the entire country and of the urgency of interventions to reduce that risk (Pagliacci and Russo, 2017).

Urgency is even more critical as most of the municipalities in the aforementioned seismic zones show weak socio-economic conditions, showing an ageing population, a lower presence of manufacturing activities and geographical remoteness. These municipalities also share tiny population figures: within seismic zone #1 and seismic zone #2, about 70% of municipalities have less than 5,000 inhabitants. Moreover, as far as the territorial distribution of these municipalities is concerned, they are mostly located in the Southern part of the country. In most cases, they represent the ridge of the Apennines, a sort of a "backbone of Italy" (as defined by Renzo Piano).

To some extent, they substantially overlap with "inner areas", a concept introduced and discussed by the Italian Nation-al Strategy for Inner Areas (Barca et al., 2014), which refers to those municipalities that – suffering from geographical remoteness and being affected by negative demographic trends – are now characterized by a significant deprivation of essential services (education, health and mobility). In their broadest definition, inner areas are defined according to a concept of spatial accessibility, as those municipalities that are located at more than 20 minutes from the nearest urban pole (i.e. a city providing the whole set of essential services under consideration). Bertolini and Pagliacci (2016) have already claimed the existence of an overlapping between inner (remote) areas and rural ones, across Italy. Furthermore, it is easy to observe an additional overlapping between inner areas and seismic zones #1 and #2. In fact, despite the presence of some major municipalities and in particular 91 major urban centres (classified as urban poles by the national strategy), the vast majority of municipalities in this portion of the country represents inner areas. (Figure 2).

In those conditions of socio-economic vulnerability, the Sendai framework for Disaster Risk Reduction 2015-2030 (UNISDR 2015) explicitly provides insights for effective actions within these seismic areas. According to it, the following dimensions should be explicitly considered when developing a plan for action at local, national and global level: i) understanding disaster risk; ii) strengthening disaster risk governance to manage dis-aster risk; iii) investing in disaster risk reduction for resilience; iv) enhancing disaster preparedness for effective response and to "Build Back Better" in recovery, rehabilitation and reconstruction.

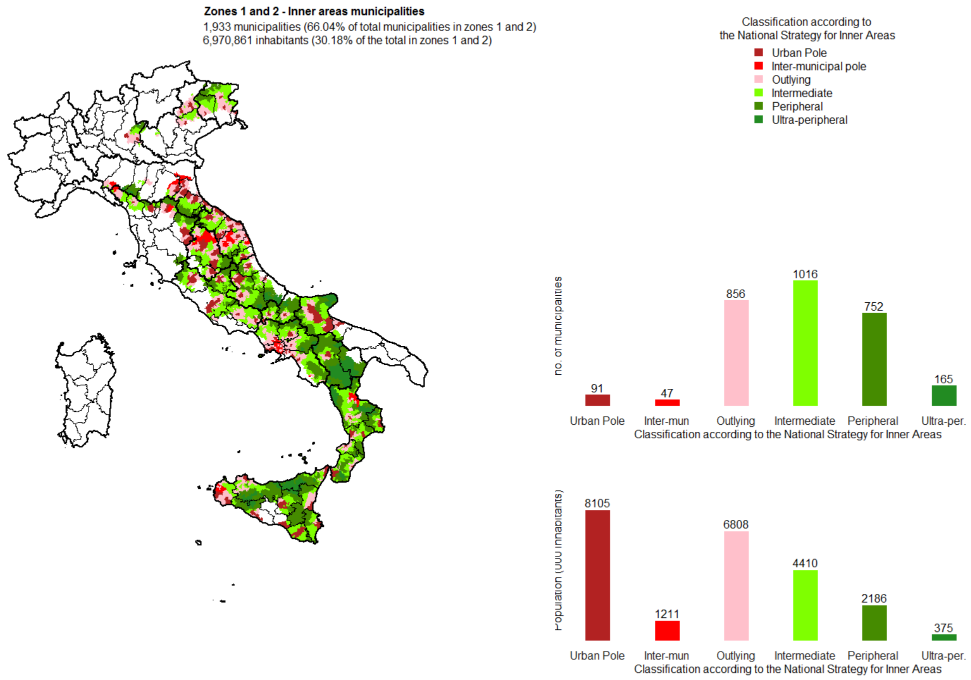
The *Piano Casa Italia*, approved by the Italian Parliament in 2016, moves in that direction. This is an ambitious long-term intervention, aimed at re-structuring private and public buildings. The Plan is conceived as a long-term intervention to bring about a significant turn in the negative effects that the present weaknesses produce in terms of human losses, not to mention the economic and social losses estimated to be several billions of euros, as potential economic damage. It is a comprehensive proposal, aimed at

[4] For each municipality, this classification returns the relative seismic hazard. On this basis, in each municipality specific anti-seismic regulations and technical norms for building construction apply (Civil Protection, 2015). Four different seismic zones (from #1 to #4) are identified: zone #1 represents the most seismic, hence dangerous, one.

[5] These figures comes from a precautionary principle, by assigning those municipalities characterized by the co-existence of more than one seismic zone to the highest seismic zone. As the only exception, there is the municipality of Rome, which comprises 19 *municipi* (boroughs), with a total population of almost 3 million inhabitants. This city has been split into two groups according to the seismic zones of its respective boroughs: in particular, 11 of them belong to the seismic zone #3, whereas the remaining 8 belong to the seismic zone #2. Moreover, an additional treatment refers to the aggregation within seismic zone #3 of the municipalities classified by the National Civil Protection into the seismic zone #3s, a typology adopted within Piedmont and Liguria (in North Western Italy). Those municipalities represent a tiny share of the total. They are just 114 municipalities, with a total population of 385,514 inhabitants, equal to 0.65% of the national population.

restructuring Italy's public buildings, houses and cultural sites, over the next decades; it has a long-term horizon that will encompass a wide territorial area with many specificities (Pagliacci et al., 2017)[6].

Figure 2 – Latest earthquakes in Italy



Source: authors' elaboration on National Strategy for Inner Areas data

The ongoing challenges of the plan are multifaceted: political, economic and social challenges pertaining to the agents asked to design the plan, to implement it and to accept it. In particular, within this broad perspective, the paper suggests the need to support informed policy measures. With regard to Italy, the territorial extension and distribution of the seismic zones makes clear the importance of focusing on a systematic analysis of socio-economic data, covering the entire country at municipality level. Although these data represent just part of the relevant matrices of information that should be used, they play a crucial role in outlining a multidimensional mapping for informing policy measures. In particular, the detection of specific heterogeneities of local areas should help policy-makers to fix place-based priorities and to define some general procedures of interventions. Although it is partial, a socio-economic analysis may significantly support the interpretation of the other layers of the aforementioned inter-connected dimensions. In particular, the aim of this paper is to return a comprehensive description of the socio-economic characteristics of the territories that the Plan should potentially target. A proper knowledge of these characteristics also represents a key issue in the critical implementation of place-based policies (Barca, 2009; Barca et al., 2012). Although there are no specific target areas, the Plan intends to start with some initial pilot interventions. By analysing a large set of socio-economic and other structural data at municipality level, the paper provides a detailed picture of seismic Italy: a contribution to ground the first priority of the Sendai framework, i.e. understanding disaster risk. It paves the way to discussing the Plan's potential impact, by considering the widest comprehensive number of Italian municipalities that could be affected by earthquake hazard, even though the Plan will address also other kinds of natural hazards (e.g., landslides, floods, etc.).

This paper takes an analytical perspective in order to explicitly support informed policy measures. Such a general aim is in line with the Sendai Framework for disaster risk reduction (UNISDR 2015). Moving from a

[6] The target of the plan is twofold: i) to set national "standard" rules for the construction of new buildings, according to anti-earthquake safety; and ii) to implement effective measures for private dwellings, and for those public buildings of strategic importance (Pagliacci et al., 2017).

detailed analysis of the socio-economic, demographic and geographic conditions across Italian territorial areas, at a municipality level (Pagliacci and Russo, 2017), this work considers a slightly different set of variables, whose statistical source is the latest round of General Censuses (years 2010-2011) and then Principal Component Analysis (PCA) and cluster analysis are adopted.

In particular, the paper is structured as follows. Section 2 provides a general overview of the adopted methodology (namely Principal Component Analysis and Cluster Analysis). Section 3 discusses available data for the analysis. Section 4 returns the main results of the analysis, which aims to distinguish groups of municipalities according to their most important structural socio-economic characteristics. Section 5 concludes the paper and outlines the major implications for the Plan and for further research.

2. Methodology

This analysis is grounded on two different methodologies, which are jointly adopted. Firstly, a Principal Component Analysis (PCA) is performed in order to highlight a fewer number of the most important components that may characterize Italian municipalities. Grounded on those extracted components, a cluster analysis is performed. Before describing the available data, subsections 2.1 and 2.2 shortly focus on each of the aforementioned methodologies.

2.1 Principal Component Analysis

Among multivariate techniques, PCA reduces the dimensions of the problem under investigation. Its first formulation is due to Hotelling (1933), although the methodology refers to Pearson (1901). PCA maximizes the amount of variance that is accounted for among the observed variables by a smaller group of variables, called Principal Components (PCs). The outcome of this kind of analysis is the reduction of the overall number of variables of a given system. In the meanwhile, PCA also preserves most of the original statistical information (Everitt and Hothorn, 2010).

In this work, PCA focuses on the major characteristics of Italian municipalities and it makes possible a reduction of the number of the most important components that characterize them. To this respect, it is used in an exploratory way, as it does not aim to model a specific factor structure. Moreover, no strong assumptions on the model itself are requested. The only ex-ante assumption made here has to do with the overall dataset under consideration. Here, PCA is adopted with the preliminary exclusion of the Italian NUTS-3 level capital cities. This choice comes from the acknowledgement that this set of municipalities is too heterogeneous if compared to the rest of the Italian municipalities.

As the original variables are set on very different scales, standardizing them avoids any distorting influences, which may come from those indicators that show higher variances. Therefore, PCA is carried out on the correlation matrix rather than the covariance matrix[7].

After having extracted the PCs, it is possible to compute the scores for the whole set of statistical units (here, for each municipality under study). In particular, on a standardized scale, each observation can be assigned a score, according to each extracted PC. In this analysis, any elaboration is performed using software R (R Core Team, 2015) and the Factominer package (Le *et al*. 2008).

2.2 Cluster Analysis

As one of the main aims of the current work is to properly identify different typologies of municipalities across Italy, cluster analysis is adopted after the extraction of PCs. Among multivariate statistical techniques, cluster analysis deals with classification issues. Indeed, it helps in identifying those units that share similar features (Tryon, 1939) and it provides a good synthesis of the structure of a dissimilarity matrix among

[7] In order to simplify the interpretation of factor loadings, principal components can be orthogonally rotated, so maintaining the un-correlation among PCs (e.g., through the VARIMAX technique). However, since rotation determines a general reduction in the explained variance, we have adopted the basic projection.

observations (Johnson, 1967). As in the case of PCA, cluster analysis is also aimed at preserving the largest part of original information[8].

From a methodological perspective, cluster analysis belongs to the unsupervised learning approaches: its outcome contributes in highlighting hidden structures within unlabeled data. In particular, through cluster analysis, a set of objects is grouped according to *p* measurable characteristics in such a way that the objects in the same group (i.e., the same cluster) are more similar to each other than to those belonging to other clusters. Thus, the analysis is based on the key concepts of 'distance' and 'similarity'

According to a chosen distance, a *n x p* data matrix is converted into a *n x n* distance matrix, a matrix containing the distances, taken pairwise, of the whole set of points. Each element d_{ij} of the matrix is then the expression of the distance between the vectors considering all the p variables. Different kinds of distance can be used[9]: here, the Euclidean distance is adopted.

As far as clustering algorithms are concerned, it has to be stressed that there are no objectively right algorithms[10]. According to this specific dataset and problem, a hierarchical cluster analysis is applied to the extracted PCs. In particular, this work adopts the *hclust* algorithm, which is provided in the R software (R Core Team, 2015). This agglomerative approach uses a set of dissimilarities (according to the Euclidean distance) for the *n* objects being clustered. The Ward's minimum variance method has been adopted in order to compute distances between pairs of clusters (Lance and Williams, 1966; Ward, 1963).

3. Data

According to this methodology, PCA and following cluster analysis are applied on a set of socio-economic variables, which essentially refer to years 2010-2011, when the latest round of general censuses took place. Five different groups of variables, available at the municipal level, are taken into account:

- Geographical features (altitude of the municipality) and characteristics of the urban settlement (population size and population density);
- Population and demographical features (foreigners, elderly people, tenancy status);
- Characteristics of buildings (age of construction and number of dwellings);
- Employment (as a whole and by economic activity);
- Agriculture and landscape.

The variables have been retrieved by the 6th General Census of Agriculture (Istat, 2010), the 15th General Census of Population and Housing (Istat, 2011a), and the 9th General Census of Industry and Services (Istat, 2011b), making possible a thorough analysis on demographic, economic and employment features of Italian municipalities. Table 1 returns each of the selected variables and their statistical source. All of them have been previously standardized.

[8] According to its properties, cluster analysis has been widely used in social sciences, in urban planning analysis and in land management issues.

[9] Such as the Minkowski distance, the Manhattan distance, the Euclidean distance. Each of them shows positive and negative issues.

[10] Clustering algorithms can be categorized into two different main typologies. While hierarchical approaches are mostly based on the core idea of building a whole hierarchy of clusters (Kaufmann and Rousseeuw, 1990), partitioning approaches divide n observations into k non-overlapping clusters. The latter group of approaches is typically based on iterative algorithms: each observation belongs to the cluster with the nearest mean (centroids). A specific objective-function can be minimized: k-means clustering (MacQueen, 1967) and k-medoids clustering techniques (Kaufman and Rousseeuw, 1990) are generally used, among partitioning approaches. Hierarchical methods do not allow the reallocation of an observation after it is assigned to a given group, while partitioning approaches are iterative. Moreover, partitioning approaches can handle larger dataset, but they do not properly managed outliers, namely observations that are numerically distant from the rest of the dataset. As far as partitioning methods are concerned, a major drawback is represented by the fact that the number of clusters to be extracted (k) has to be set in advance. In fact, this number is empirically identified.

Table 1 – List of variables for PCA

	Variable	*Definition*	*Label*	*Source*
Geography/ urban settlement	Altitude	Altitude of the town hall, in meters above sea level	altitudine_del_centro_metri	15th Census Population Housing
	Population size	Total Population	pop_11	
	Population density	Resident people per km^2	densita_11	
Population and demography	Old Age people	Ratio of older dependents (people aged 65+) to the total population	perc_65anni_11	15th Census Population and Housing (2011)
	Foreigners	Ratio of the foreign population to the total population	perc_stranieri	
	Tenancy	Ratio of the families living as tenants in their dwellings, to the total number of families living in dwellings	perc_affitto	
Buildigns	Pre-1949 buildings	Buildings built before 1949 over the total number of buildings	ed_pre2GM	15th Census Population and Housing (2011)
	Buildings with 3 dwellings and more	Share of buildings with at least three dwellings	ed_almeno3int_perc	
Employment	Employment	Number of employees (all economic activities)	addetti_tot	9th Census Industry and Services (2011)
	Wholesale and retail trade	Ratio of the employees in manufacturing local units (Nace Rev.2) to the total number of employees	addetti_c_perc	
	Manufacturing	Ratio of the employees in wholesale and retail trade; repair of motor vehicles and motorcycles local units (Nace Rev.2) to the total number of employees	addetti_g_perc	
Agriculture and Landscape	share of UAA	UAA (Utilised agricultural area) over total land area (%)	perc_sau	6th Census Agriculture (2010)
	Woodland per inhabitants	Woodlands (in hectare) per resident people	boschi_abit	
	Average size of agricultural units	Average size of agricultural units (in hectare)	dim_media_uagri	

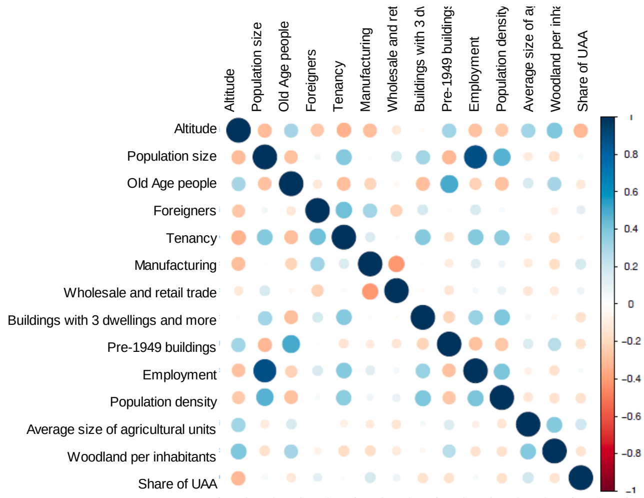
Source: authors' elaboration

The aforementioned set of variables have been retrieved for all the Italian municipalities. As Census data are mostly used, municipalities in 2011 are considered for the analysis (8,092 across the country). However, as already stressed, this analysis does not take into account the Italian NUTS-3 level capital cities (*capoluoghi di provincia*), for they simply are too different from the rest of the Italian municipalities. Given the exclusion of these 110 observations, the final dataset comprises 7,982 observations.

Before performing PCA, the correlation matrix has bee analyzed. In

Figure 3, it can be observed that, as expected, the largest positive correlation is between population size and employment. Moreover, before performing PCA, suitability of variables for PCA has been verified, by means of the Kaiser–Meyer–Olkin (KMO) test[11]. This test suggests that the variables under study can be considered as satisfactory (KMO=0.697).

Figure 3 – Correlation matrix



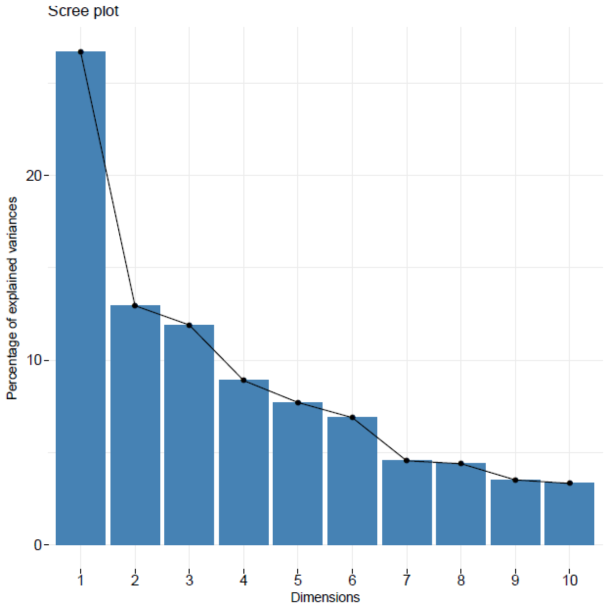
Source: authors' elaboration

[11] This is a sampling adequacy test, which may range from 0 to 1 (Kaiser, 1974).

4. Results: PCA and Cluster analysis

As a preliminary analysis of the percentage of explained variance of each of the extracted PCs (just 10 PCs are considered here), the Guttman–Kaiser criterion[12] is applied to our data: eight PCs are considered here, accounting for 75% of total variance (see Figure 4).

Figure 4 – PC extraction (explained variance)



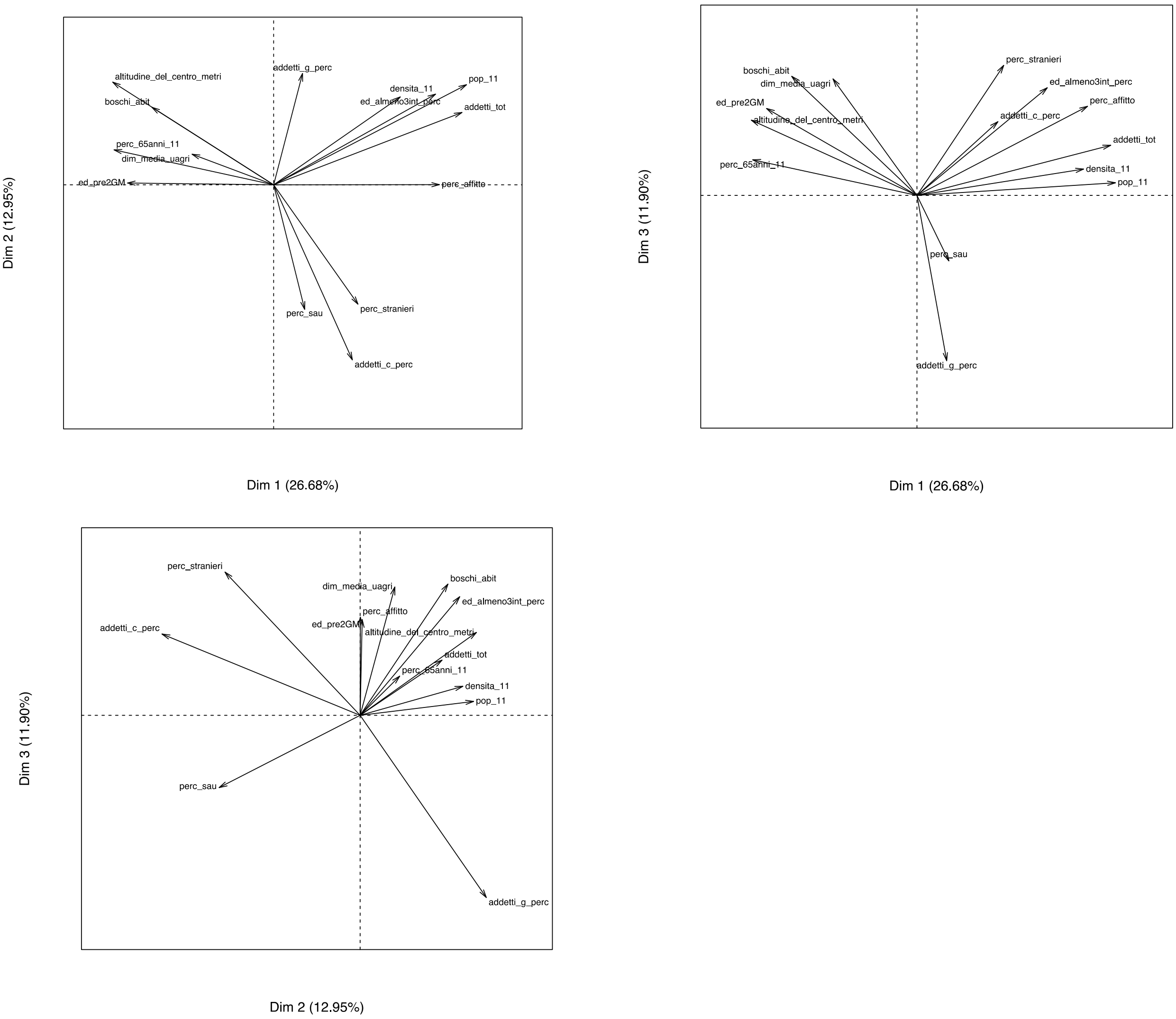
Source: authors' elaboration

In order to interpret each of the extracted PCs (or dimensions), relations between original variables and the eight extracted PCs have to be considered. For the sake of brevity, Figure 5 just shows the combination of the first three dimensions. By analyzing the plots, it is easy to observe that PC1 (Dim.1) is positively linked to total population and total employment, but also to share of families living as tenants and to the share of foreigners out of the total. Conversely, it is negatively related to the altitude of the municipality. Thus, PC1 seems to be a proxy of large metropolitan areas, located in the flatlands. PC2 is positively related to the share of employment in the retail sector and to the altitude of the municipality. Conversely, it is negatively related to the share of employment in manufacturing out of the total and the share of foreigners. PC3, again, is negatively related to the share of employment in the retail sector, while it is positively related to the share of foreigners.

In more general terms, the extracted eight components confirm the importance of the following characteristics: geographical remoteness vs. urban centrality, different structures of the economy (e.g. presence of manufacturing activities vs. share of tertiary employment out of the total), characters of the rural areas, other socio-demographic features (e.g. ageing population or tenancy status).

[12] In this way, we can choose those PCs explaining at least 70–80% of the cumulative variance.

Figure 5 – PC extraction: different factor spaces

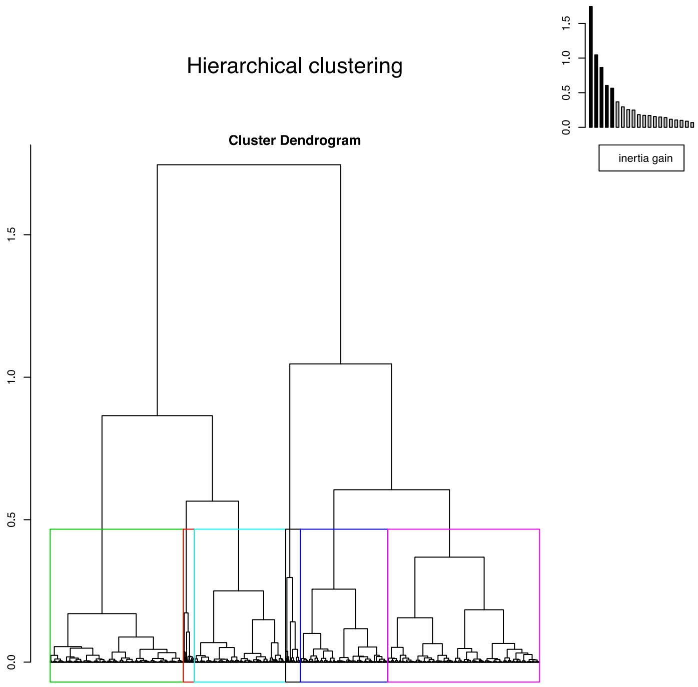


Source: authors' elaboration

As already stated in Section 2, these dimensions represent the input variables that make possible to perform the cluster analysis on the total number of Italian municipalities. When considering the first eight PCs, the distances between all the observations have been computed. Based on that, a hierarchical agglomerative cluster analysis has been performed, by considering the Ward's linkage as a criterion of agglomeration of clusters. Thanks to the observation of the inertia gain, it seems plausible to to cut the dendrogram in order to retrieve six different clusters, as shown in Figure 6. Each cluster is characterized by a very different number of observations: the two smallest ones comprise 290 and 331 municipalities, respectively. On the opposite, the other six clusters include more than 1,500 municipalities each. The largest one includes 2,036 municipalities.

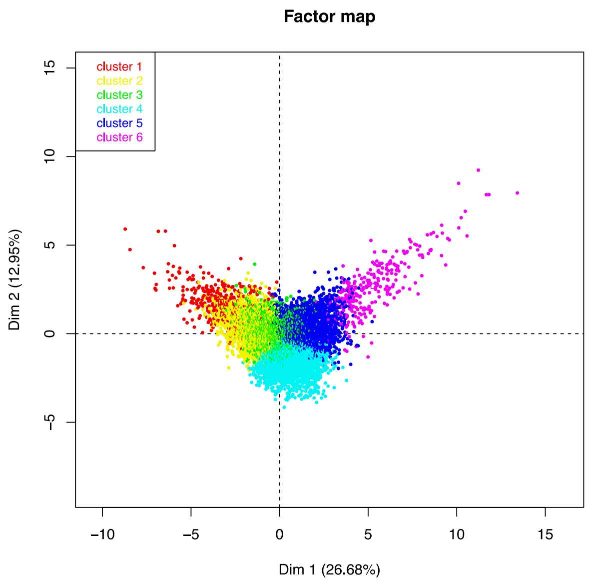
The extracted clusters also differ in terms of their own features. While Figure 7 just gives an idea of the way municipalities are grouped into clusters, when considering the two-most-significant-PCs Cartesian space, Table 2 returns a more comprehensive description of each of the extracted clusters, by considering those variables originally adopted in the PCA.

Figure 6 –Cluster analysis: the dendrogam



Source: authors' elaboration

Figure 7 – Cluster analysis: a representation of the outcome on the first two PCs map



Source: authors' elaboration

According to the average characteristics, it is possible to label each cluster in the following way:

- Cluster 1 (red; 331 municipalities): least-densely populated municipalities mainly located in mountain areas (963 meters above sea level, on average), with no manufacturing activities (10% of total employment) and a small share of foreigners (4% of the total). In this cluster, more than 50% of buildings were built before WWII, while the share of elderly people is the highest (28.2% of total population);
- Cluster 2 (yellow; 1036 municipalities): small municipalities mainly located in hill areas (577 meters above sea level), with a limited presence of manufacturing activities (17.7% of total employment).

This cluster show many similarities with cluster 1, for instance in terms of presence of elderly people (26.9% of the total), old buildings and limited presence of buildings with three dwellings or more;

- Cluster 3 (green; 1886 municipalities): medium municipalities with no foreigners (2.8% of the total) and a huge share of tertiary activities (retail accounts for 28.8% of total employment). This cluster shows additional economic weakness if considering total employment, which is definitely lower than the one observed across similar clusters;
- Cluster 4 (light blue; 1897 municipalities): manufacturing municipalities (47.6% of total employment), with a large share of foreigners (8.9%). This cluster also shows higher population density and a lower share of elderly people out of the total;
- Cluster 5 (blue; 1542 municipalities): medium-large towns of about 10,000 inhabitants, with foreign people (accounting for 7.1% of the total);
- Cluster 6 (purple; 290 municipalities): large towns (37,000 inhabitants on average). This cluster shows socio-economic features that in most cases are similar to the ones observed across the NUTS-3 level capital cities, such as a lower share of elderly people, a large number of modern buildings, with more than three dwellings each.

Table 2 – Clusters' composition and descriptive statistics

	cluster 1	cluster 2	cluster 3	cluster 4	cluster 5	cluster 6
altitudine_del_centro_metri	963.4	577.9	328.4	191.1	234.3	123.3
pop_11	823.0	1326.8	3643.2	3857.5	9126.2	37497.1
perc_65anni_11	0.2824	0.2694	0.2222	0.2086	0.1922	0.1927
perc_stranieri	4.2619	4.9845	2.8290	8.8810	7.0924	6.3570
perc_affitto	8.1388	9.1913	8.6642	13.2923	16.5643	20.3700
addetti_c_perc	10.1755	17.7735	16.0260	47.5653	24.5500	23.7598
addetti_g_perc	14.4678	17.8944	28.8477	14.1448	23.3647	24.9218
ed_almeno3int_perc	25.0463	16.2248	11.9411	18.3722	33.2850	39.2987
ed_pre2gm	52.2444	51.9533	24.7763	31.1945	21.9819	17.4478
addetti_tot	245.0	226.2	505.8	1208.6	2302.7	10373.1
densita_11	13.3	60.5	131.5	207.9	630.5	1839.1
dim_media_uagri	59.3696	8.8672	8.1933	10.8570	5.8134	6.5137
boschi_abit	2.3217	0.3952	0.1225	0.0676	0.0466	0.0082
perc_sau	3256.08	2341.49	5020.94	5533.29	3092.13	3301.74
No. of municipalities	331	2036	1886	1897	1542	290

Source: authors' elaboration

The most interesting finding of this analysis refers to the comparison of clusters' composition and clusters' features on one hand and the presence of the four seismic zones, as defined by the Civil Protection (2015), on the other. To this respect, Table 3 returns the number of municipalities within each seismic area, by considering both the country level and distinguishing each single cluster. Results are particular insightful. At national level, municipalities within the most seismic zone (#1) are just 8.7%. However, they are 17.1% of the total number of municipalities of the cluster 3 (municipalities with no foreigners) and 12.7% of the total number of municipalities of the cluster 1 (municipalities in mountain areas). Larger towns are definitely under-represented when considering seismic zone 1. When considering seismic zone 2, results are partially different. Indeed, while 35% of municipalities of cluster 3 are in this class, even 39.3% of large town (cluster 6) and 35.9% of NUTS-3 level capital cities are included into this seismic zone.

Table 3 – Clusters' composition, per seismic zone

	Cluster	Zone 1		Zone 2		Zone 3		Zone 4		Total
		No.	%	No.	%	No.	%	No.	%	
1	Small municipalities (mountains)	42	12.7	76	23.0	129	39.0	84	25.4	331
2	Small municipalities (hills)	228	11.2	550	27.0	682	33.5	576	28.3	2036
3	Medium municipalities no foreigners	323	17.1	661	35.0	267	14.2	635	33.7	1886
4	Manufacturing muncipaliteis	42	2.2	390	20.6	983	51.8	482	25.4	1897
5	Medium-large towns (foreginers)	57	3.7	385	25.0	695	45.1	405	26.3	1542
6	Large towns	6	2.1	114	39.3	107	36.9	63	21.7	290
	NUTS-3 level capital ciites	7	5.5	46	35.9	49	38.3	26	20.3	128
	Total	**705**	**8.7**	**2222**	**27.4**	**2912**	**35.9**	**2271**	**28.0**	**8110**

Source: authors' elaboration

5. Conclusions

The quantitative analysis presented in this work aims at pointing out the major differences that characterise Italian municipalities. These differences will play a key role when considering policy implementation at country level. In particular, the recently approved Piano Casa Italia, as a comprehensive proposal aimed at restructuring Italian public buildings and houses over the next decades, will be confronted not only with the impressive magnitude of the interventions needed, but also with its very long-term horizon and with these huge territorial differences.

The detailed analysis proposed in this paper provides brand-new knowledge about those specific socio-economic differences, in particular considering the interactions between four domains, which look as complementary: after shock emergency, recovery of areas hit by recent earthquakes, risk reduction, socio-economic development. These domains are coordinated respectively by the National Civil Protection, *Struttura Commissariale*, the Piano Casa Italia, the National Strategy for Inner Areas.

As a main result of this analysis, it largely confirms some of the findings that had been preliminarily outlined by Pagliacci and Russo (2017)[13]. In particular, the fact that, in Italy, seismic hazard tend to couples with socio-economic vulnerability. Indeed, the territorial allocation of the clusters that show the greatest social and economic vulnerability tends to overlap with the most seismic zones. This is the case, for instance, of the municipalities across the mountain areas of the South of the country.

It would be important that the Piano Casa Italia would address some specific policy interventions in order to take these criticalities into account.

First of all, geography and demography significantly matter. Most seismic zones in Italy are characterized by lower population density and population size than the rest of the country. Most of these municipalities belong to mountain areas (especially throughout the Apennines, across Southern Italy). These features could make the implementation of the first steps of the Plan more difficult, both in terms of available competences and high-skilled workers at local level, because of a generalized lack of larger cities. In addition, population size and distribution by age are important (Pagliacci and Russo, 2017). Demographic features of seismic zones #1 and #2 (and in particular a larger share of elderly people) are expected to increase the vulnerability of these regions, making a rapid implementation of the Piano Casa Italia even more important.

Secondly, local economies of seismic municipalities show greater weakness than the economy in the rest of the country. Less people employed in the manufacturing could also turn into a smaller amount of local wealth, which could have a negative impact on the implementation of the Plan, itself.

[13] Thus, most of the following conclusive considerations are drawn from that previous work (Pagliacci and Russo, 2017).

Thirdly, age and type of buildings matter. Older buildings (i.e. the ones that were constructed before WWII) in seismic zones #1 and #2 represent the primary target of the Plan. These buildings deserve immediate attention for a restructuring that would also take into account anti-seismic regulations. As claimed by the Plan itself, in its political intentions, a crucial step in the Plan is the assessment of the actual state of preservation of buildings especially in seismic zones #1 and #2. Clusters with municipalities of these zones tend to be characterised by buildings that comprise a few dwellings each. This aspect could play a positive role when considering the total amount of time needed to assess the building's current state, but a fragmented ownership may slow down the overall process of interventions (Pagliacci and Russo, 2017). On the opposite, appropriate economic incentives could enhance the speed of such a process by fostering local interventions by a network of professional experts. Here, clear and transparent procedures will play a major role in ensuring responsibility and ethical behaviour to reduce the medium-long term tragic effects experienced in many past disasters in Italy. The Plan will both support the bottom-up ability of local communities to innovate and promote higher building standards across Italy. Its success will critically depend on ethical and legal training and will also be dependent on other changes in the Property Register and in the digitalisation of the procedures for building permits (Pagliacci and Russo, 2017).

Although all the above-mentioned items constitute relevant indications in designing the Plan, there is an aspect that deserves special attention: governance of the Plan at national and local level. The number of municipalities involved will require coordination of all the many specific features and peculiarities of the municipalities in question. To not address this issue could become the Achilles' heel of the Plan. Considering the Plan only as a top-down process might not be effective, but also considering it as a bottom-up procedure it might be hard to produce effective results, given the lack of local competences and the need for adopting the most effective technical standards and best-practice techniques.

Discussing this issue is the next step in the ongoing research that we are carrying out to support a collective discussion on the Plan and possibly a more effective decision-making process to implement it. Another strand of complementary action-research would be needed to assess the vulnerability of physical infrastructures as well as of social infrastructures (health services, social services, education, communication), and of public management at a local level, all of which have effects at a local level in enhancing preparedness. This action-research must be a pillar for making the local process effective in understanding and mastering the long-term implications of the Plan.

References

Barca, F. (2009), "An Agenda for the Reformed Cohesion Policy. A Place-based approach to meeting European Union challenges and expectations", Report to the Commissioner for Regional Policy, Brussels, April.

Barca, F., McCann, P., & Rodríguez–Pose, A. (2012) The Case for Regional Development Intervention: Place-Based versus Place-Neutral Approaches, *Journal of Regional Science*, 52(1), pp. 134-152.

Barca, F., Casavola P. and Lucatelli S. (eds.) (2014) *A strategy for inner areas in Italy: definition, bjectives, tools and governance*. Materiali Uval Series 31.

Barone, G., Benni F., Brasili C., & Mocetti S. (2013) Una stima degli effetti economici di breve periodo del terremoto in Emilia-Romagna. *Politica Economica*, 2, 199–214.

Bertolini, P. and Pagliacci, F. (2016) Territorial unbalances in quality of life. A focus on Italian inner and rural areas. DEMB Working Paper Series, 87.

Chubb, J. (2002) Three earthquakes: Political response, reconstruction, and the institutions. In J. Dickie, J. Foot, F. M. Snowden (eds.), *Disastro!: Disasters in Italy since 1860: Culture, politics, society*. New York. NY: Palgrave.

Civil Protection (2015) Seismic classification. Available at http://www.protezionecivile.gov.it/jcms/en/classificazione.wp (latest access: 11 Jan. 2017).

Everitt, B.S., Hothorn, T. (2010) *A Handbook of Statistical Analysis using R*. Boca Raton (FL): Taylor & Francis Group, 2nd Ed.

Guidoboni, E. (2012) *Il peso economico e sociale dei disastri sismici in Italia negli ultimi 150 anni*. Bologna: Bononia University Press.

Guidoboni, E. (2014) *L'Italia dei disastri. Dati e riflessioni sull'impatto degli eventi naturali 1861-2013*. Bologna: Bononia University Press.

Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417–441. doi:10.1037/h0071325

Istat (2010) 6° Censimento Generale dell'Agricoltura. http://dati.istat.it.

Istat (2011°) 15° Censimento Generale della Popolazione e delle Abitazioni. http://dati.istat.it.

Istat (2011b) 9° Censimento Generale dell'Industria e dei Servizi. http://dati.istat.it.

Johnson, S.C. (1967) Hierarchical clustering Schemes. *Psychometrika*, 32 (3): 241-254.

Kaiser, H.F. (1974) An index of factorial simplicity. *Psychometrika* 39: 31-36.

Kaufman, L., Rousseeuw, P. (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken (N.J.): Wiley Series in Probability and Mathematical Statistics.

Lance, G.N.; Williams, W.T. (1966) A Generalized Sorting Strategy for Computer Classifications. *Nature* 212(5058): 218.

Le. S., Josse. J., Husson. F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1), 1-18. Doi:10.18637/jss.v025.i01.

MacQueen, B. (1967) Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley (CA): University of California Press, vol. 1, pp. 281-297.

Pagliacci, F., Russo, M. (2017) Earthquake hazard in Italy. Cluster analysis of socio-economic data to inform place-based policy measures, DEMB Working Paper Series n. 110.

Pagliacci, F., Russo, M., Sartori, L. (2017) Social Innovation and Natural Disasters: the "Casa Italia" Plan, *Sociologia Urbana e Rurale*.

Pearson, K. (1901) On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series* 6, 2, 559–572. doi:10.1080/14786440109462720

R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Tryon, R. C. (1939) *Cluster analysis*. New York: McGraw-Hill.

UNISDR (United Nations International Strategy for Disaster Reduction) (2015) Sendai Framework for Disaster Risk Reduction, 2015–2030. Geneva: United Nations.

Ward, J.H. (1963) Hierarchical Grouping to Optimize an Objective Function. *Journal of American Statistical Association* 58: 236-244.