DEMB Working Paper Series

N. 82

Socio-economic effects of an earthquake: does sub-regional counterfactual sampling matter in estimates? An empirical test on the 2012 Emilia-Romagna earthquake

Francesco Pagliacci*, Margherita Russo*

April 2016

* University of Modena and Reggio Emilia
CAPP (Center for the Analysis of Public Policies)
Address: Viale Berengario 51, 41121 Modena, Italy
email: margherita.russo@unimore.it
email: francesco.pagliacci@unimore.it

ISSN: 2281-440X online

Socio-economic effects of an earthquake: does sub-regional counterfactual sampling matter in estimates?
An empirical test on the 2012 Emilia-Romagna earthquake

Francesco Pagliacci and Margherita Russo

Dipartimento di Economia Marco Biagi and CAPP (Centro Analisi Politiche Pubbliche), Università degli Studi di Modena e Reggio Emilia, viale Berengario 51, 41121 Modena, Italy
Email: francesco.pagliacci@unimore.it; margherita.russo@unimore.it

Paper accepted for the 56th ERSA Congress "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria, Special Session: *Counterfactual methods for regional policy evaluation*

Abstract

Estimates of macroeconomic effects of natural disaster have a long tradition in economic literature (Albala-Bertrand, 1993a; 1993b; Tol and Leek, 1999; Chang and Okuyama, 2004; Benson and Clay, 2004; Strömberg, 2007; UNISDR, 2009; Cuaresma, 2009; Cavallo and Noy, 2009; Cavallo *et al.*, 2010; The United Nations and The World Bank, 2010). After the seminal contribution of Abadie *et al.* (2010) in identifying synthetic control groups, with DuPont and Noy (2015) a new strand has been opened in estimating long term effects of natural disaster at a sub-regional scale, at which the Japan case provides plenty of significant economic variables. Although the same methodology has been applied in estimating the impact of earthquakes in Italy (Barone *et al.* 2013; Barone and Mocetti, 2014), the analysis has been limited to the regional scale. In our paper, due to a lack in long-term time series data at municipality level, this paper cannot adopt the methodology suggested by Abadie *et al.* (2010). Nevertheless, it provides a test bed for assessing the relevance of a sub-regional counterfactual evaluation of a natural disaster's impact.

By taking the 2012 Emilia-Romagna earthquake as a case study, we propose a comprehensive framework to answer some critical questions arising in such analysis. Firstly, we address the problem of identifying the proper boundaries of the area affected by an earthquake. Secondly, through a cluster analysis we show the importance of intra area differences in terms of their socio-economic features. Thirdly, counterfactual analysis is assessed by adopting a pre- and post-earthquake difference-in-difference comparison of average data in clusters within and outside the affected area. Moreover, three frames to apply propensity score matching at municipality level are also adopted, by taking the control group of municipalities (outside the affected area): (a) within the same cluster, (b) within the same region, (c) in the whole country. The four variables considered in the counterfactual analysis are: total population; foreigner population; total employment in manufacturing local units; employment in small and medium-sized manufacturing local units (0 to 49 employees).

All the counterfactual tests largely show a similar result: socio-economic effects are heterogeneous across the affected area, where some clusters of municipalities perform better, in terms of increase of population and employment after the earthquake, against some others. This result sharply contrasts with the average results we observe by comparing the whole affected area with the non-affected one or with the entire region.

Keywords: cluster analysis, counterfactual analysis, Emilia-Romagna, earthquake
JEL codes: C38, R11, R58

This paper has been realized in the research project Energie Sisma Emilia (www.energie.unimore.it) of the University of Modena and Reggio Emilia (Italy). The project is co-funded by Fondazione Cassa di Risparmio di Modena, Azienda Unità Sanitaria Locale di Modena e Università di Modena e Reggio Emilia.

1. Introduction

In May 2012, a severe earthquake hit a wide area of Northern Italy, lying across Emilia-Romagna, Veneto and Lombardy. The earthquake left a huge scar on a densely populated and wealthy area: in particular, largest part of damages occurred to Emilia-Romagna municipalities. With a population of around 550,000 inhabitants, they represent one of the country's most productive areas, producing almost 2% of the national GDP and significantly contributing to regional (and national) exports. Furthermore, the same area has been the object of several socio-economic studies, due to its peculiarities, such as: (i) its important industrial and agri-food districts, which firstly originated in the 1960s and 1970s (Brusco, 1989) and (ii) its local governance, which is the result of a special balance between public and private action[1]. Thus, in addition to human losses, the earthquake caused large damages to both material infrastructures and immaterial socio-cultural components. The former mainly concerned houses (around 31,000 were left unfit for living), historical and cultural buildings, commercial and industrial structures, buildings for health and social services. The latter mostly refer to the immaterial fractures to local social and cultural system (businesses, public administrations, families and their networks), whose re-composition is not easy to predict.

Moving from the large extent of damages within this specific case study, this work aims to contribute to the more general issue of modelling the assessment of socio-economic effects of a natural disaster. Here, we suggest the application of a counterfactual analysis (i.e. a comparison between the affected area and a similar, but non-affected, one) to properly quantify the economic effects of an earthquake, in both the short and medium term.

Nonetheless, from a regional analysis perspective, this approach poses some critical questions. A first issue deals with the proper definition of the affected area. Such a definition has always represented a key task in earthquakes effects studies (see, for instance, Centro di Portici, 1981, on the Irpinia earthquake). Actually, affected area's boundaries are likely to be blurred, although policymakers have tried to strictly define

[1] The so-called "Emilia model", which has recently returned to being a topic of discussion after the publication of some Italian essays (cf. De Maria, 2012; Mosconi, 2012).

them (at a municipality level), in order to implement policies and to define beneficiaries to be admitted to funding schemes. Here, the task is crucial, in particular in the framing of the counterfactual analysis at a sub-regional level. To this respect, territorial units we adopt here are single municipalities.

A second issue deals with the intuition that the affected area itself does not represent a homogeneous region, especially in its socio-economic features. The 2012 earthquake hit a wide region, whose single parts show different characteristics: actually, both manufacturing municipalities and more agricultural ones coexist within the boundaries of the affected area. Moreover, tiny rural towns occur together with medium-sized urban poles. Thus, all these characteristics might play a role in assessing a counterfactual analysis of the earthquake effects. Thus, this paper provides a cluster analysis, which covers all municipalities throughout Emilia-Romagna, to detect groups of similar municipalities.

Eventually, moving from the cluster analysis results, a counterfactual 'difference-in-differences' analysis is undertaken: in particular, pre- and post- earthquake situation is compared among affected and non-affected municipalities (treatment and control group, respectively). This analysis is based on the major changes that affected population and employment trends.

This work is structured as follows. Section 2 focusses on the theoretical background of the analysis of the effects of an earthquake and some answers that have been given in order to identify the territorial unit of analysis. Section 3 underlines the main methodologies that have been adopted in order to define the affected areas and carry on a counterfactual analysis. Section 4 returns some major results from both cluster and four sets of counterfactual analyses. Section 5 suggests some policy conclusions and further development of the research.

2. Theoretical background

Studies on natural hazards (Kahn, 2003; Barone and Mocetti, 2014) suggest that different aspects drive the effects produced by an earthquake on a given area (e.g. location of the hypocenter, distance from the epicentre, morphological characteristics of the affected area). Furthermore, economic studies have also focuses on the socio-economic effects that an earthquake (as well as other natural disasters) might produce. This large strand of research (Albala-Bertrand, 1993a; 1993b; Tol and Leek, 1999; Chang and Okuyama, 2004; Benson and Clay, 2004; Strömberg, 2007; UNISDR, 2009; Cuaresma, 2009; Cavallo and Noy, 2009; Cavallo *et al.*, 2010; The United Nations and The World Bank, 2010) points out the fact that human activities and other socio-economic characteristics might affect both the amount of damages and the speed and efficiency of the reconstruction process.

DuPont and Noy (2015) explicitly observe medium and long terms effects of an earthquake. They suggest that the occurrence of a seismic event might influence overall path of growth of the affected area (e.g. a country or a region). To this respect, conclusions diverge when considering either the short or the medium to long term. Although in the long run natural disasters hardly manage to divert an entire country from its natural growth path (Fiske, 2012), in the shorter term, economic conclusions differ. The path of growth, especially at a local level, might be diverted, although the effects could be positive in some cases, because of the large amount of public resources that, being produced externally, flow into a limited area to support the reconstruction process. Other contributions focus on the differences between natural disasters' impacts in developed and developing countries. Cross-country approaches tend to claim that the latter suffer the worst effects due to: i) their lower capacity to increment public investment; ii) a limited organization in the emergency phase; iii) a lower degree of prevention of natural disasters (Albala-Bertrand, 1993a; 1993b; Strömberg, 2007; The United Nations and The World Bank, 2010). Neither the presence of democratic institutions increases the adverse effects produced by natural hazards (Kahn, 2003). Rather, when controlling for the amount of resources, some studies consider the quality of public spending (namely, allocated resources actually spent on the territory) as a fundamental driver for triggering a positive process of reconstruction (e.g., modernizing fabric and physical capital) (Barone and Mocetti, 2014).

Nonetheless, the use of national data to estimate the effects of an earthquake can provide some biases. Like a flood or a landslide, its effects are spatially concentrated: showing a reduced significance at wider scales, any national analysis underestimates its possible socio-economic effects. To control for these biases, DuPont and Noy (2015) studied the Kobe earthquake of 1995 and its effects on a narrower region. They compared economic trends in both the prefecture hit by that earthquake and a so-called 'pseudo-Kobe' (i.e. a comparable prefecture that was not affected by that earthquake), by suggesting a counterfactual model. Also Barone and Mocetti (2014) analyzed the effects of an earthquake at sub-national level. Taking into consideration two of the largest earthquake in Italy since WWII (namely, the Friuli 1976 earthquake and the Irpinia 1980 one), they adopted the methodology of the synthetic control group, suggested by Abadie *et al.* (2010), showing a divergence in long-term effects. Focusing on the 2012 Emilia-Romagna earthquake, Barone *et al.* (2013) quantified the economic impact of that earthquake as rather limited. They got this result, by considering the regional economy as a whole and comparing it to other similar regions across Italy. Nevertheless, despite some methodological innovations[2], they did not explicitly tackle the issue of defining the most appropriate unit of analysis. Regional analyses, although being more appropriate than national ones, still underestimate the earthquakes' effects: indeed, affected municipalities just represent 20% out of the region, in terms of both land area and population (Pagliacci and Bertolini, 2015). Barone *et al.* (2013) neither answered to the second question this kind of analysis poses, namely the identification of most appropriate control group, within which selecting counterfactual examples.

To this respect, there is a wide room to implement quantitative analysis to refine previous works on this topic. In particular, the following sections will suggest some methodological tools to answer the aforementioned questions.

3. Data and methodology

3.1. Identifying the affected area

Firstly, this paper aims to properly identify the area affected by the 2012 Emilia-Romagna earthquake. In 2012, policymakers explicitly debated this issue, as they were asked to identify the exact boundaries of the affected area where funds for private reconstruction could be allocated (Ranuzzini *et al.*, 2015). Thus, rather than from a

[2] They did not take into account GPD growth: rather, they referred to a regional indicator produced by RegiosS on infra-annual basis (Benni and Brasili, 2006).

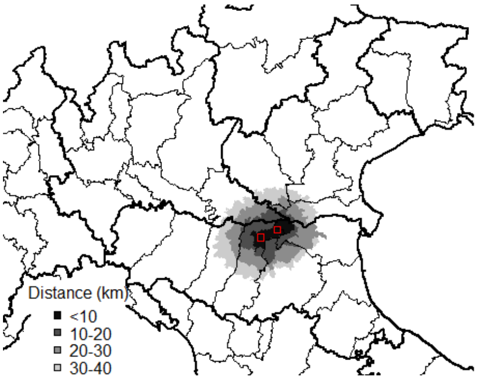
geographical perspective (i.e. according to a given distance from the epicentre[3]), the boundaries of the affected area have been mostly identified according to the several decrees that were issued during the emergency phase[4]. In particular, three decrees have adopted 'technical' criteria (drawn up in collaboration with the national department of Civil Protection) to define those municipalities[5]: referring to the alternative lists mentioned by them, a certain amount of arbitrariness occurs when defining the boundaries of the affected areas. In our analysis we refer to the list provided by the D.L. no. 74/2012, with the exclusion of the municipality of Ferrara. Thus, in the following, we will always refer to 32 municipalities as the ones affected by the earthquake. Indeed, other lists refer to a too wide area: for instance, the 58-municipalities list includes the four affected NUTS 3 level capital cities. Figure 1 and Figure 2 compare a geographical definition of the affected area, according to the distance of each municipality from the closest epicentre, with the aforementioned institutional lists of affected municipalities. Focusing on the institutional definitions, Table 1 summarises the socio-economic features of the affected area, referring to latest available year. When moving from 32-municipalities to 58-municipalities area, population largely increases, because of the inclusion of the four NUTS 3 level capital cities of the region. These results suggest that including them into the affected area would distort the analysis (they lie far apart from the epicentres, and they have actually suffered limited damages). Conversely, the identification of 32 most affected municipalities appears to be the most appropriate to consider as the area with the bigger damages.

[3] In fact, two distinct tremors occurred, each of them with its own epicentres.

[4] Piazzi *et al.* (2015) suggest alternative methodologies to identify the area affected by the earthquake, at both municipality (the number of times the orders by the Commissioner Delegate cite each municipality) and supra-municipality level (referring to affected local market areas, namely *sistemi locali del lavoro* in Italian). Nonetheless, both these choices are unfeasible, as they introduce distortions into outcomes.

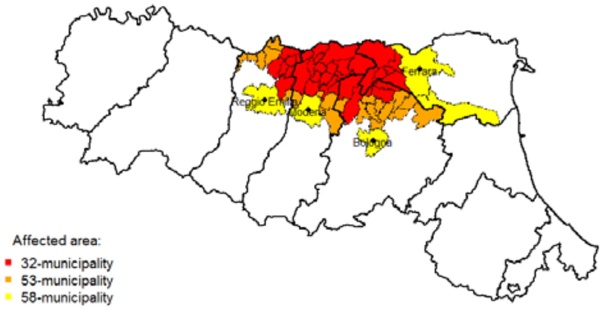
[5] The three decrees providing the lists of affected municipalities are: Decree Law (D.L.) no. 74, of 6 June 2012, identified 33 municipalities in Emilia-Romagna for which the Civil Protection had certified the presence of structural damages to urban fabric and other buildings; Ministerial Decree (D.M.) of MEF (the Italian Ministry of economy and finances), of 1st June 2012, identified 53 municipalities in Emilia-Romagna for which suspension of tax compliance and other tax benefits were made available; Ordinance no. 29, of 28 August 2012, identified the main criteria for allocating grants for the immediate restoration of buildings and other residential units that had been damaged by the earthquake in 58 municipalities.

Figure 1 – Municipalities affected by the earthquake: distance from the epicentres (in red)



Source: authors' elaboration

Figure 2 – Municipalities affected by the earthquake, according to different decrees



Source: authors' elaboration

Table 1 – Socio-economic features of the affected area, according to the 32-municipality, 53-municipality and 58-municipality lists (municipality-level average values, year 2014)

	D.L. 74/ 2012 *) 32-municipalities	D.M. 1 2012 53-municipalities	Ordinance 29/2012 58-municipalities
Total population	415,426	631,435	1,528,487
Population 0-14years	60,877	93,848	206,122
Population 65+ years	90,121	134,244	354,319
Number of families	172,263	262,861	704,614
Total foreigners	53,317	77,784	208,221
EU foreigners	8,266	12,646	34,575
Non-EU foreigners	45,051	65,138	173,646
Ageing index	148	143	172
Population 65+ years (as a % out of total population)	21.7%	21.3%	23.2%
Population 15-64 years (as a % out of total population)	63.7%	63.9%	63.3%
Birth rate (per a thousand persons)	9.2	9.2	8.7
Foreigners as a % out of total population	12.8	12.3	13.6
Population density	233.83	233.88	384.88
Total land area (square kilometres)	1,777	2,700	3,971
Artificial surfaces (square kilometres)	199	310	583
Continuous urban fabric surfaces	72	105	201
Discontinuous urban fabric surfaces	39	56	90
Industrial and Commerical surfaces	53	84	147
Roads (in kilometres)	4,659	7,482	12,799
Agricultural holdings	8,402	12,131	16,640
Total Agricultural Area (in hectares)	141,908	212,285	298,127
Utilised Agricultural Area (in hectare)	129,210	191,838	269,451
Agricultural holdings employees	24,118	34,339	48,802
Active firms	30,979	46,878	133,250
Active firm empolyees	120,180	194,331	578,819
Local units	33,490	50,970	144,432
Local unit employees	128,452	206,252	556,319
Local units – manufacturing (C)	6,401	9,126	15,621
Loca units – construction (F)	4,847	7,543	16,494
Local units – wholesale and retail trade, transportation and storage, accomodation and food service activities (G + H + I)	10,732	16,999	45,212
Local units – other services	11,510	17,302	67,105
Local unit employees – manufacturing (C)	58,968	92,407	154,648
Local unit employees – construction (F)	12,419	18,761	40,856
Local unit employees – wholesale and retail trade, transportation and storage, accomodation and food service activities (G + H + I)	31,654	56,210	172,332
Local unit employees – other services	25,411	38,874	188,483

Source: authors' elaboration on Emilia-Romagna Region data (StRia)

3.2. Cluster analysis and available data

After having defined the boundaries of the affected 32 municipalities, a Cluster Analysis (CA) makes possible the detection of some typologies of municipalities which characterise that area[6]. Here, we adopt a hierarchical approach, as it is more suitable for properly handling outliers and it does not require any *ex ante* definitions of the number of clusters. Furthermore, its outcome can be graphically displayed through a bi-dimensional diagram, i.e. a dendrogram (Kaufmann and Rousseeuw, 1990). Eventually, we compute the dissimilarity matrix by applying the Euclidean distance, whereas the Ward's method is chosen to compute distances between clusters. Accordingly, at each stage, the pair of clusters with minimum between-cluster distance are merged and total within-cluster variance is minimised[7] (Lance and Williams, 1966; Ward, 1963).

According to this methodology, CA is here performed on a set of socio-economic variables, which refer to pre-earthquake conditions. Reference period is years 2010-2011, when the latest round of general censuses took place. Actually, variables have been retrieved by 6th General Census of Agriculture (Istat, 2010), 15th General Census of Population and Housing (Istat, 2011a), 9th General Census of Industry and Services (Istat, 2011b), making possible a thorough analysis on demographic, economic and employment features of Emilia-Romagna municipalities. Table 2 shows selected variables and their statistical source. All of them have been previously standardised and some of them refer to the share out of the total at regional level[8].

Table 2 – Input variables of CA, source and reference year

Label of the Variable	Description	Source	Year
For_pop Tot_popt	Foreign population Total population	15th General Census of Population and Housing	2011
Area TAA_Area (%)	Land area (in square kilometres) Total agricultural area and land area ratio (%)	6° General Census of Agriculture	2010
Manuf_employm SMEs_manuf_employm	Employees in manufacturing local units (Nace Rev.2) Employees in small and medium manufacturing local units (0 to 49 employees)	9th General Census of Industry and Services	2011

Source: authors' elaboration

[6] By means of CA, observations are grouped in such a way that the units in the same cluster are more similar to each other than to those belonging to other groups (Kaufman and Rousseeuw, 1990). According to a chosen distance, we convert a n x p data matrix into a n x n distance matrix, which contains the distances, taken pairwise, of a set of points, each element of the matrix d_{ij} being the expression of the distance between the vectors, considering all the p variables.

[7] CA is performed with Stata 12.

[8] They represent single municipality's contribution to the regional level.

Although data is available for all municipalities across Emilia-Romagna, in our CA we have preliminarily excluded ten NUTS 3 level capital cities (*capoluoghi di provincia*), in order to get accurate results. These capital cities, none of them occurring within the boundaries of the affected areas, differ from other municipalities in terms of socio-economic features. Being the identification of a counterfactual sample for those municipalities affected by the 2012 earthquake the main aim of this analysis, they have not been considered: 338 municipalities compose the final set of observations.

For each of the aforementioned variables, Table 3 returns main descriptive statistics, for the affected (32) municipalities and non-affected (306) ones. Despite major differences in the number of observations, the affected area is characterised by a larger share of foreigners and of manufacturing activities out of the total than non-affected one[9].

Table 3 – Descriptive statistics: average values across different classifications

Variables	**Affected area** 32 obs.	**Non-affected area (excluding NUS3 level capital cities)** 306 obs.	**Emilia-Romagna (excluding NUTS 3 level capital cities)** 338 obs.	**NUTS 3 level capital cities** 10 obs.	**Emilia-Romagna** 348 obs.
Total population	12,718.38	7,542.69	8,032.70	162,708.40	12,477.40
Foreigners	1,484.47	711.49	784.67	18,681.60	1,298.95
Land area (ha.)	5,551.93	5,905.31	5,871.86	26,059.02	6,451.95
Total agricultural area (ha.)	4,434.62	3,479.28	3,569.72	17,096.48	3,958.42
Employees in manufacturing	1,818.72	906.49	992.86	11,245.10	1,287.46
Employees in manufacturing SMEs	1,091.59	505.45	560.95	6,176.50	722.31
Total population (as a % out of the region)	0.29	0.17	0.18	3.75	0.29
Foreigners resident (as a % out of total of the municipality)	11.90	8.95	9.23	11.31	9.29
Land area (as a % out of the region)	0.25	0.26	0.26	1.16	0.29
Rurality Index (ratio between total agricultural and land area)	0.79	0.61	0.62	0.62	0.62
Employees in manufacturing (as a % out of total employees of the municipality)	50.47	34.31	35.84	18.49	35.34
Employees in manufacturing SMEs (as a % out of total manufacturing employees of the municipality)	62.44	70.61	69.83	57.35	69.48

Source: authors' elaboration on data of 6th General Census of Agriculture (Istat, 2010), 15th General Census of Population and Housing (Istat, 2011a), 9th General Census of Industry and Services (Istat, 2011b).

3.3. Counterfactual analysis

Counterfactual analysis makes possible the comparison between what actually happened and what would have happened in the absence of a given intervention. Thus, it evaluates those specific changes that can be attributed to an intended (or unintended)

[9] Pearson correlation coefficients among variables and their territorial distribution at municipality level are listed in Piazzi *et al.* (2015).

intervention (cause-effect questions). Given its properties, counterfactual analysis usually applies to public policies evaluation: it allows researchers and policymakers to make judgements on specific interventions, based on the observation of quantitative data. Such analysis has been undertaken here in order to assess the effects of the 2012 Emilia earthquake: the main question under investigation is the comparison between what happened to municipalities affected by it in terms of population and employment trends and what would have happened in its absence.

From a theoretical perspective, the effect of a given intervention (i.e. a 'treatment') refers to the overall variation in the outcome that is explicitly driven by it ($\Delta Y|\Delta X$). Nonetheless, the lack of treatment to be compared with cannot be observed in reality. Thus, Campbell e Stanley (1966) suggested quasi-experimental design frameworks to analyse counterfactual effects. A population under study can be split into a 'treated' group and a 'control' one: if two groups are statistically equivalent, observed counterfactual effects will represent a good proxy of the real ones. To get statically equivalent groups, experimental designs adopt randomized control trials, whereas quasi-experimental designs cannot do (selection bias might arise). To avoid it, several techniques can be applied. Unfortunately, the synthetic control method, suggested by Abadie *et al.* (2010) and insightfully adopted by DuPont and Noy (2015) with regard to the Kobe earthquake, cannot be implemented here. Indeed, no long-term time series data are available at municipal level. In particular, we have only six census data before the earthquake (1971-2011) and as a result, poor fits of the synthetic control method are returned, when comparing treated observations to control ones[10].

Given the available data, propensity score matching (PSM) has been applied, as it is more suitable for the case under study here. PSM uses statistical models to compute the probability of 'being treated', based on a set of observable characteristics. Eventually, it matches treated and non-treated observations with similar probability scores (i.e. with the most similar characteristics, before the event[11]) (Rosenbaum and Rubin, 1983). PSM defines a control group after the treatment, making possible a

[10] The authors have tried to compute synthetic control method (Abadie *et al.*, 2010) through the R package "Synth" (Abadie *et al.*, 2011) for comparing affected municipalities with non-affected municipalities across Emilia-Romagna. Results are available upon request.

[11] Probability (or propensity) score is just a proxy of the distance between two observations in a population.

statistical balance among groups, according to observed characteristics[12] and assuming that no distortions arise from omitted variables.

The first step in undertaking PSM analysis refers to the identification of most similar control group's units to be compared with the treated group. Such an identification comes from the *propensity score*, namely the conditional probability that a unit, given its characteristics, will be assigned to the treatment itself. This probability is computed by means of a logistic regression model, where the treatment represents the binary dependent variable.

After having estimated the propensity score for each treated and non-treated observation, most similar observations are matched. Among the matching techniques to be implemented to detect the most similar observations, here the nearest neighbour matching procedure is adopted: it matches each observation from the control group to just one treated observation, based on closeness of their propensity scores (Rosenbaum and Rubin, 1983; Olmos and Govindasamy, 2015).

After having identified most similar participants, difference-in-differences (DID) methods are adopted: by using data collected at both baseline and end-line for treatment and control groups, they get rid of selection bias, under the assumption that unobservable factors determining selection are time invariant. Actually, DID methods simultaneously compare both pre-post situations and 'treated'-'non treated' observations. In the former case, counterfactual example just refers to the *ex-ante* situation in the same municipality (this assumption lies on the idea that the variable under study would not have changed, if the event did not occur)[13]. In the latter one, counterfactual example refers to the comparison between treated observations and control ones, after the event. Here, a major assumption lies in the fact that starting conditions were equal: thus the importance of PSM in selecting the control group and the idea of controlling observations by cluster (which would probably reduce significant

[12] Compared to parametric regression, PSM does not impose a given shape on the relationship between regressors and dependent variable. Indeed, it can be both linear and non linear.

[13] A further assumption lies in the idea that no spontaneous dynamics are observed among variables. This seems to be justified by the fact that a short time span is under study here (from December 2011 to December 2012, the earthquake having occurred in May 2012. Nonetheless, if that holds true for some variables (e.g. demographic ones), employment variables might be affected by short terms economic trends.

differences in starting conditions). When jointly considering two effects, it is possible to assess properly the overall effect of a given intervention on a given area in a given time span.

According to this general theoretical framework, the case under study here shows some peculiarities. As already stressed, we are interested in defining the effects of an unintended natural event, having affected a number of municipalities with different socio-economic features. Thus, a problem of pseudo-randomization occurs, although being not perfect. According to Tobler's First Law of Geography (Tobler, 1970), both distance to the epicentres and – in general – neighbourhood do matter in characterising municipalities. Furthermore, even socio-economic features do not occur evenly across municipalities: spatial proximity matters. Thus, being randomized control trial unfeasible here, this paper adopts a quasi-experimental counterfactual framework. Firstly, population and employment trends have been analysed at municipality level (before and after the earthquake), by considering the following variables: i) total population; ii) foreigner population; iii) total employment in manufacturing local units; iv) employment in SMEs manufacturing local units (0 to 49 employees). Variations of population variables refer to the period Jan. 2012-Jan. 2014; variation of employment variables refer to the period Oct. 2011 (census round) to 2013 (average annual value).

In each case, the treated group is defined as the set of 32 affected municipalities (as specified in section 3.1). Four alternative sets of control groups are selected according to the following specifications:

- PSM is not applied: counterfactual examples are built according to each extracted cluster, by comparing affected and non-affected municipalities in each cluster within the affected areas and in the corresponding clusters outside the affected area[14];
- PSM is applied to each Emilia-Romagna municipality, but the 32-municipalities list, by controlling for cluster (as specified in section 3.2); thus, each municipality is linked to the most similar one among those that belongs to the same cluster;
- PSM is applied to each Emilia-Romagna municipality, but the 32-municipalities list;

[14] Those clusters that do not include any of the 32 municipality list have not been considered.

- PSM is applied to each municipality in the country.

After having specified a control group, DID is applied as aforementioned specified. Eventually, all results are returned by disentangling clusters: this choice makes results' interpretation clearer and more accurate[15].

4. Results

4.1. Dendrogram and cluster labelling

The output of the hierarchical CA is graphically shown in Figure 3 through a dendrogram. The optimal number of clusters to be selected can be detected by maximising the Calinski–Harabasz Index (Caliński e Harabasz, 1974), which is based on the variance between groups (Figure 4). The index suggests the best classification for a two–cluster classification that roughly grained separates the most manufacturing municipalities from all the others. When moving to a three-cluster classification, the smallest municipalities (mostly located in the Apennines) are grouped, and a further disaggregation in four clusters, reaching a local optimum, separates the largest towns. Any further disaggregation implies a decrease in the index until the next local optimum is reached with a ten-cluster classification. This result is due to the separation of the three largest towns under observation. In our analysis we consider 10 clusters which provide groups more adequately differentiated than a four-clusters disaggregation.

Figure 3 - Dendrogram

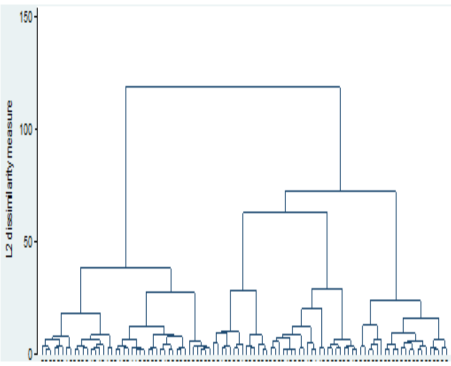
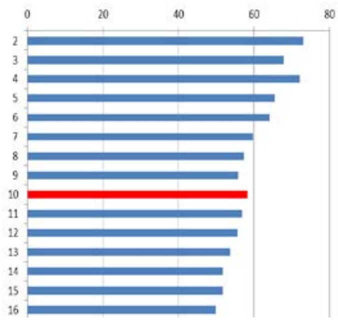


Figure 4 - Calinski-Harabasz Index (2 to 16 partitions)

Number of partitions	Calinski-Harabasz pseudo-F
2	73.09
3	67.78
4	72.11
5	65.58
6	64.09
7	59.63
8	57.4
9	55.8
10	58.25
11	56.84
12	55.62
13	53.63
14	51.74
15	51.73
16	49.85



Source: Piazzi *et al.* (2015)

[15] Counterfactual analysis and PSM have been performed by means of software R (R Core Team, 2015) and package Matching (Sekhon, 2011).

Table 4 – Clusters' composition: number of municipalities (affected and non-affected area)

	Cluster										Total
	1	2	3	4	5	6	7	8	9	10	
Non-affected area	2	14	18	58	41	51	46	26	20	30	306
Affected area	1	1	6	8	6	10	0	0	0	0	32
Total	3	15	24	66	47	61	46	26	20	30	338

Source: authors' elaboration

From Table 4 we observe that some clusters have no municipalities affected by the earthquake. Table 5 returns the main characteristics for each cluster, by referring to those variables adopted in CA[16]. Similar information is also provided for the group of NUTS 3 level capital cities (which had been previously excluded from the analysis). In order to properly describe (and label) each cluster, the map in Figure 5 shows their territorial distribution throughout Emilia-Romagna and highlights the municipality and the NUTS 3 level boundaries as well as the 32-municipality area affected by the 2012 earthquake. By observing average values per clusters, two dimensions seem driving the overall output, although clusters have been actually obtained by referring to a set of six variables. They are: i) population size and ii) share of manufacturing employment. In particular, when ordering clusters according to their population size, from the largest to the smallest one, and referring population size and share of employment in manufacturing activities to the respective averages calculated on Emilia-Romagna region, without NUTS 3 level capital cities (*capoluoghi di provincia*), the clusters may be labelled as follows:

- Cluster #1 – Larger towns: it just includes three municipalities, which also show lower shares of manufacturing activities out of the total and a very high rurality index (because of their large total municipality area);
- Cluster #2 – Larger towns, with a very low presence of manufacturing activities: the economy of these municipalities is mostly centred on services; they also show large population density;
- Cluster #3 – Large towns with a high rurality index: although being populous municipalities, they are characterised by a large presence of agricultural areas (as measured by the rurality index);

[16] For further details on the characteristics of each cluster, refer to Piazzi *et al.* (2015).

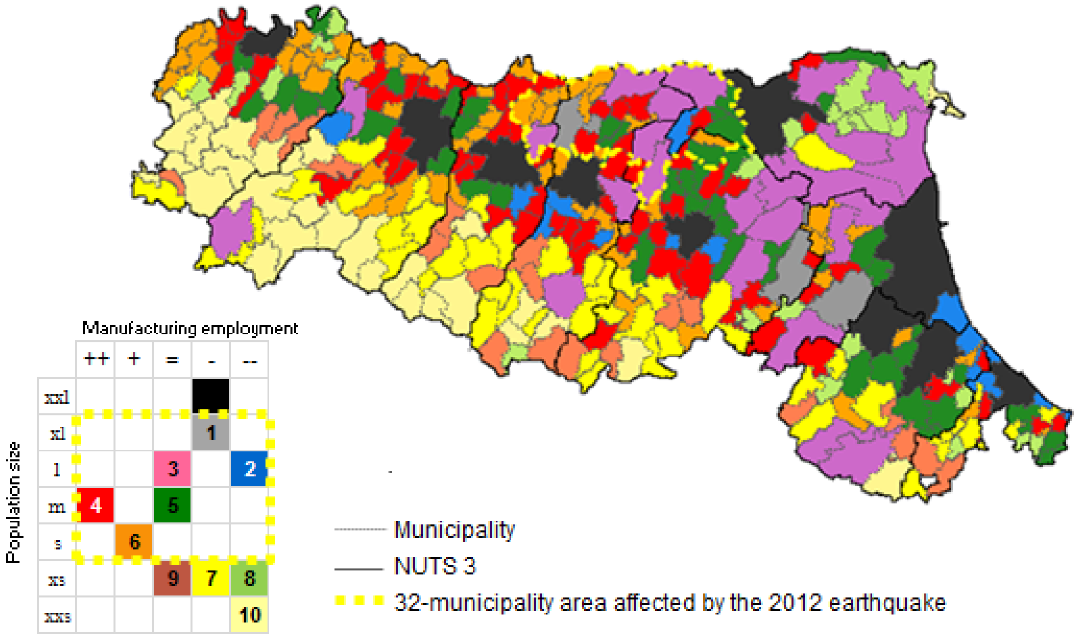
- Cluster #4 – Medium-sized towns, with a large presence of manufacturing activities: in this cluster, large enterprises mostly characterise the manufacturing sector;
- Cluster #5 – Medium-sized towns, with a high rurality index: compared to cluster #4, this group of municipalities shows a similar population size, although it is characterised by a larger share of agricultural areas;
- Cluster #6 – Small towns, with a large share of foreigners out of total population: the municipalities within this group also have some manufacturing large enterprises;
- Cluster #7 – Smaller towns: these municipalities, showing the presence of some manufacturing activities, are mostly located in the hills;
- Cluster #8 – Smaller towns, with a limited amount of SME manufacturing activities;
- Cluster #9 – Smaller town, with a very high rurality index;
- Cluster #10 – Smallest towns: they are mostly located in the mountain area.

Table 5 – Clusters' composition and descriptive statistics (average values per cluster and in Emilia-Romagna)

	clusters										**Emilia-Romagna (excluding capital cities)**	**NUTS 3 level Capital cities**	**Emilia-Romagna**
	1	2	3	4	5	6	7	8	9	10			
	Larger towns (3)	Larger towns, very low manufacturing activities (15)	Large towns, high rurality index (24)	Medium-sized towns, with manufacturing activities (66)	Medium-sized towns, very high rurality index (47)	Small towns, largest presence of foreigners (61)	Smaller towns (20)	Smaller towns, with a limited amount of SMEs (46)	Smaller towns, very high rurality index (26)	Smallest towns (20)			
Total population	64302.7	27035.0	17155.0	8181.3	7830.3	6882.1	3725.0	3345.9	2928.5	2006.0	**8032.7**	162708.4	12477.4
Foreigners	6911.7	2594.3	1589.4	753.6	619.1	935.5	365.5	243.1	237.0	121.6	**784.7**	18681.6	1299.0
Land area (ha.)	18410.8	4026.5	15311.9	4115.5	4680.6	3497.1	6322.0	3808.7	6570.6	9179.7	**5871.9**	26059.0	6451.9
Total agricultural area (ha.)	14169.8	2233.8	10880.5	2813.9	3561.0	2420.9	2920.2	2980.0	2877.7	3309.9	**3569.7**	17096.5	3958.4
Employment in manufacturing activities	7491.0	1975.4	1773.6	1825.0	983.2	880.0	241.0	100.6	262.1	54.4	**992.9**	11245.1	1287.5
Employment in manufacturing SMEs	4391.3	1291.5	973.0	818.3	670.5	517.1	215.1	100.6	125.1	54.4	**560.9**	6176.5	722.3
Population (as a % out of total region)	1.481	0.623	0.395	0.188	0.180	0.158	0.086	0.077	0.067	0.046	**0.185**	3.747	0.287
Foreigners (as a % out of population)	10.749	9.761	8.890	9.305	8.083	13.656	9.026	6.781	7.549	5.278	**9.231**	11.314	9.290
Land area (as a % out of total region)	0.820	0.179	0.682	0.183	0.208	0.156	0.282	0.170	0.293	0.409	**0.262**	1.161	0.287
Rurality Index (ratio of total agricultural area and land area)	0.761	0.492	0.734	0.696	0.768	0.686	0.447	0.763	0.438	0.358	**0.624**	0.616	0.624
Employment in manufacturing activities (as a % out of total employment)	33.441	20.221	36.749	54.152	38.965	40.881	29.723	17.451	36.270	12.727	**35.837**	18.494	35.339
Employment in manufacturing SMEs (as a % out of manufacturing employment)	57.830	76.087	58.070	44.814	72.577	60.074	94.707	100.000	45.275	100.000	**69.834**	57.348	69.475

Source: authors' elaboration on Istat data

Figure 5 – Territorial distribution of clusters and of ten NUTS 3 level capital cities



Legenda (in parentheses, number of municipalities per cluster)

Largest cities
- NUTS 3 level Capital cities (10)
- 1 Larger towns (3)
- 2 Larger towns, very low manufacturing activities (15)
- 3 Large towns, high rurality index (24)

Medium-sized and small towns
- 4 Medium-sized towns, with manufacturing activities (66)
- 5 Medium-sized towns, very high rurality index (47)
- 6 Small towns, largest presence of foreigners (61)

Smallest towns
- 7 Smaller towns (20)
- 8 Smaller towns, with a limited amount of SMEs (46)
- 9 Smaller towns, very high rurality index (26)
- 10 Smallest towns (20)

For clusters 1 to 10: population size and share of employment in manufacturing activities refer to the respective averages calculated on Emilia-Romagna region, without NUTS 3 level capital cities (*capoluoghi di provincia*).
Yellow border highlights those clusters that comprise at least one municipality affected by the 2012 earthquake.
Source: authors' elaboration

As shown in Figure 5, the area affected by the 2012 earthquake mostly comprises clusters of medium-sized and large towns. Conversely, referring to the major economic activity, that area shares both manufacturing traits (clusters #4 and cluster #6) and more agricultural/rural traits (especially among municipalities in cluster #3 and #5). Nevertheless, whereas manufacturing activities mostly occur in the Western portion of that area, the easternmost portion of that area shows a more rural and agricultural economy. Such a distinction confirms the idea that the area that comprises the aforementioned 32 municipalities cannot be considered as a unit of analysis as it were a homogeneous territory.

To summarize, cluster analysis does not just provide insightful details about territorial characteristics of the affected area, it also provides some suggestions to detect the control group to be compared with.

4.2. Estimating counterfactual effects of the 2012 earthquake

In order to estimate the effects of the 2012 earthquake, we consider two population variables (total population and foreigners) and two employment variables (in total manufacturing and in manufacturing SMEs). By comparing the aggregate effect of the earthquake in the affected area, the non-affected one, and in the Emilia-Romagna region (Table 6), we observe that, in the two-years' time span under study (i.e. year 2012-2014), total population in the region and in the non-affected area has grown more than in the affected area. In the same period foreigners increase their presence more outside the affected area (signalling a weakening of the demand of labour in that area). With regard to manufacturing employment, its decrease has been generalized, even though with heavier effects in the affected area but the decrease in employment in SMEs has been lower[17].

Table 6 – Pre- and post-earthquake socioeconomic variation in the affected area, the non-affected one, and in the Emilia-Romagna region

	Total Population	Foreigners	Employment in manufacturing	Employment in manufacturing SMEs
Affected Area	+1.71%	+10.84%	-3.65%	-3.05%
Non-affected area	+2.50%	+18.24%	-1.49%	-4.98%
Emilia-Romagna	+2.42%	+17.46%	-1.77%	-4.71%

Source: authors' elaboration on Istat data

To take into account the different socio-economic characteristics of the municipalities within the affected area, in the following sub-sections we present the results of four sets of counterfactual analysis. Firstly, the identified clusters are adopted to perform a pre- and post-earthquake difference-in-difference comparison of average data in clusters within and outside the affected area. Secondly, we adopt three frames to apply propensity score matching at municipality level, by taking the control group of municipalities (outside the affected area): within the same cluster (a), within the same region (b), in all the country (c).

[17] A more detailed analysis of the effects in the ten clusters is presented in Piazzi *et al.* (2105).

The four variables considered in the counterfactual analysis are: total population; foreigner population; total employment in manufacturing local units; employment in small and medium-sized manufacturing local units (0 to 49 employees).

Cluster by cluster DID

We apply DID to each cluster, by disentangling affected and non-affected municipalities[18]. Table 7 returns the effects that are observed when comparing affected and non-affected municipalities on total population, foreigners, total manufacturing employment and manufacturing employment in SMEs. These results return interesting evidence: affected municipalities in most populous clusters #1 and #2 (namely, the towns of Carpi and Cento, respectively) show a larger population and employment increase than the non-affected ones. Many factors can explain this outcome. Firstly, the entity of damages has been relatively lower in Carpi and Cento than in neighbouring municipalities. Secondly, the increase in population can be explained by their larger availability of unoccupied (and not damaged) properties. Thirdly, the increase of employment in manufacturing can be explained by the substituting effect of supply in those municipalities with respect to similar specialisation offered by the neighbours, more heavily affected.

Conversely, affected municipalities in clusters #4, #5, #6 (i.e. those groups of medium-sized towns, showing more manufacturing traits than the regional average) have been heavily affected by the 2012 earthquake. They have experienced a worse performance than non-affected ones when considering almost all variables. In particular, the group of most manufacturing affected municipalities (cluster #4) have experienced the worst performance, compared with non-affected ones. Moreover, the affected municipalities in cluster #6, characterised by the largest share of foreigners in 2011, have registered a plunge in the presence of foreigners in years 2012-2013, compared to non-affected municipalities belonging to the same cluster.

[18] Cluster that do not include any of the 32-municipalities list (namely, clusters #7, #8, #9 and #10) have not been considered. Furthermore, besides 10 NUTS 3 level capital cities, additional observations have been excluded from this analysis. Indeed, due to fusion processes occurring among municipalities, 2012 and 2013 data are not available for Massa Fiscaglia, Migliarino, Migliaro, Poggio Berni, Torriana, Sissa, Trecasali, Bazzano, Castello di Serravalle, Crespellano, Monteveglio, Savigno.

Municipalities in cluster #3 (namely large towns with high rurality index) have experienced a decrease in manufacturing employment after the earthquake, although their population as well as foreign people have increased within the same period of time. To this respect, observed trend here is rather similar to that observed in Cento and Carpi.

Table 7 – Counterfactual effects, DID, per variable and per cluster

	1	2	3	4	5	6
Total population	623.5	577.9	106.9	-120.4	31.0	-95.0
Foreigners	299.5	271.7	84.8	-9.2	-9.8	-54.6
Employment in manufacturing	209.6	166.0	-57.0	-144.6	2.9	-43.6
Employment in manufacturing SMEs	104.0	107.1	34.0	-40.9	-13.3	3.6

Figures are estimated with DID, by comparing pre-post average values between affected and non-affected municipalities within each cluster.
Total population and foreigners: effects are estimated by comparing data on 1st Jan. 2012 and on 1st Jan. 2014.
Employment: effects are estimated by comparing data in 2011 Census (Oct. 2011) and average values for year 2013.
Source: authors' elaboration on Istat data

Nevertheless, the results shown in Table 7 may suffer from some distortions, arising from the fact that a different number of observations is included into treatment and control groups (because of clusters composition). In particular, it could be misleading to refer to the whole clusters when undertaking a counterfactual analysis.

P-score: municipalities in the Emilia-Romagna Region and in Italy

PSM is applied to this analysis, in order to link each affected municipality to the most similar non-affected municipality. Firstly, a logit model has been run on 326 observations[19]. Table 8 returns model main results (model 1). All variables but land area are statistically significant. Thus, the same model has been re-estimated by removing that non-significant regressor (model 2). Eventually, propensity scores for each of the municipality under study have been estimated through model 2.

According to the results from model 2, matching between treated and control observations has been obtained by selecting nearest neighbour matching, allowing for replacement of observations and choosing a one-to-one matching. Three alternative ways to detect the nearest neighbours have been explored: with a selection: (a) within each cluster; (b) among all observations throughout Emilia-Romagna; (c) among all

[19] Details on this sample are in Footnote 18.

observations in the whole country. Table 9 shows matched observations in those three cases.

Table 8 – Logit regression: main results

	Model 1			Model 2		
	Estimate		Std. error	Estimate		Std. error
Population (as a % out of total region)	2,609	**	0.913	2,175	**	0.636
Foreigners (as a % out of population)	0.165	*	0.072	0.171	*	0.071
Land area (as a % out of total region)	-1,178		1,714	-		-
Rurality Index (ratio of total agricultural area and land area)	0.130	**	0.033	0.126	**	0.031
Employment in manufacturing activities (as a % out of total employment)	0.110	**	0.026	0.113	**	0.025
Employment in manufacturing SMEs (as a % out of manufacturing employment)	0.033	*	0.015	0.034	*	0.015
Intercept	-21,330		3,798	-21,408	**	3,764

*,** statistically significant at 5% and 1%, respectively.
Number of observations: 326
Source: authors' elaboration on Istat data

In order to test the robustness of matching when not considering a same-cluster control observation, we present in Table 10 estimated counterfactual effects, obtained by means of DID and then aggregated per cluster. When computing PSM on observations belonging to the same clusters (a), the most affected clusters are #4 and #6, where a stronger decrease in both population and employment variables has occurred since 2012. Conversely, larger towns (namely, cluster #1 and #2) have experienced a positive performance, compared to non-affected municipalities within the same clusters. Similar results are returned, when computing PSM on all observations throughout Emilia-Romagna (b) and on all observations in the whole country (c). In the case (b), cluster #3 performance is less negative, whereas cluster #5 performance worsens. With regard to the comparison to the whole country results confirm what observed in the regional comparison. A finer analysis of the counterfactual control group should be detailed, even though a first sight on the selected observations confirms the robustness of this result.

Table 9 – Definition of the control group (PSM, nearest neighbour with replacement): PSM on (a) all observations and on (b) observations belonging to the same cluster (number of cluster is returned)

Affected municipality	cluster	Control Municipality (belonging to the same cluster) (a)	cluster	Control Municipality (Emilia-Romagna) (b)	cluster	Control Municipality (Italy) (c)	Nuts3
Carpi	1	Imola	1	Luzzara	6	Altamura	BA
Cento	2	Formigine	2	Argenta	3	Filottrano	AN
Bondeno	3	Castelfranco Emilia	3	Gattatico	5	Foligno	PG
Correggio	3	Castelfranco Emilia	3	Faenza	1	Ceneselli	RO
Crevalcore	3	Castelfranco Emilia	3	Villanova sull'Arda	6	Ca' d'Andrea	CR
Finale Emilia	3	Budrio	3	Budrio	3	Meduna di Livenza	TV
Mirandola	3	Castelfranco Emilia	3	Imola	1	Asolo	TV
San Giovanni in Persiceto	3	Copparo	3	Copparo	3	Trapani	TP
Cavezzo	4	Mordano	4	Ziano Piacentino	6	Oldenico	VC
Medolla	4	San Martino in Rio	4	San Martino in Rio	4	Bevilacqua	VR
Pieve di Cento	4	Pianoro	4	Lesignano de Bagni	7	Novi Ligure	AL
Ravarino	4	Sant Agata Bolognese	4	Sant Agata Bolognese	4	Zimella	VR
Reggiolo	4	Mordano	4	Poviglio	5	Atessa	CH
San Felice sul Panaro	4	Sant Agata Bolognese	4	Sant Agata Bolognese	4	Castelvisconti	CR
Sant Agostino	4	Mordano	4	Ziano Piacentino	6	Tolentino	MC
Soliera	4	Soragna	4	Castelfranco Emilia	3	Atessa	CH
Bomporto	5	Gattatico	5	Cortemaggiore	6	Scurelle	TN
Mirabello	5	Gattatico	5	San Martino in Rio	4	Calendasco	PC
Poggio Renatico	5	Russi	5	Fidenza	3	Bra	CN
San Pietro in Casale	5	Meldola	5	Meldola	5	Sondalo	SO
San Prospero	5	Poviglio	5	Portomaggiore	7	Saluggia	VC
Vigarano Mainarda	5	Montegridolfo	5	Molinella	3	Nereto	TE
Campagnola Emilia	6	Luzzara	6	Luzzara	6	Castiglione delle Stiviere	MN
Camposanto	6	Villanova sull'Arda	6	Villanova sull'Arda	6	Fossato di Vico	PG
Concordia sulla Secchia	6	Spilamberto	6	Nonantola	4	Arzignano	VI
Fabbrico	6	Villanova sull'Arda	6	San Martino in Rio	4	Roverchiara	VR
Galliera	6	Pontenure	6	Fiorano Modenese	4	Abbiategrasso	MI
Novellara	6	Luzzara	6	Luzzara	6	Occimiano	AL
Novi di Modena	6	Luzzara	6	Sant Agata Bolognese	4	Sorgà	VR
Rio Saliceto	6	Cadeo	6	Castelfranco Emilia	3	Montichiari	BS
Rolo	6	Villanova sull'Arda	6	Calendasco	4	Villanova sull'Arda	PC
San Possidonio	6	Cortemaggiore	6	Cortemaggiore	6	San Cipriano Po	PV

Source: authors' elaboration on Istat data

Table 10 – Counterfactual effects: results per variable and per cluster

		1	2	3	4	5	6
(a)	Total population	222.0	603.0	-352.3	-101.3	81.7	-60.2
	Foreigners	40.0	463.0	-81.3	11.5	-7.8	-24.1
	Employment in manufacturing	578.7	185.6	-159.0	-181.6	23.4	26.8
	Employment in manufacturing SMEs	216.0	205.6	-75.6	-57.4	24.2	33.7
(b)	Total population	2046.0	1091.0	-228.8	-197.0	-134.0	-220.0
	Foreigners	1376.0	485.0	-122.3	-47.1	-130.3	-69.3
	Employment in manufacturing	209.6	141.9	38.6	-207.6	77.4	-20.5
	Employment in manufacturing SMEs	64.6	47.0	88.4	-62.2	-22.1	19.7
(c)	Total population	1589,0	1150,0	99,8	-108,6	11,3	-348,8
	Foreigners	1000,0	693,0	47,0	-57,6	-66,8	-187,6
	Employment in manufacturing	261,9	-22,6	76,5	-40,0	13,9	300,0
	Employment in manufacturing SMEs	203,5	36,9	80,3	-80,3	-69,8	-1,9

Total population and foreigners: effects are estimated by comparing data on 1st Jan. 2012 and on 1st Jan. 2014. Employment: effects are estimated by comparing data in 2011 Census (Oct. 2011) and average values for year 2013.
Source: authors' elaboration on Istat data

5. Conclusions

The earthquake of 2012 in Emilia-Romagna hit a wide area, with large socio-economic differences within the affected area itself. Because of these features, assessing an appropriate counterfactual analysis is not an easy task. Firstly, this paper provides a specific contribution in defining the boundaries of the area affected by the earthquake. Despite alternative definitions, this work mostly refers to institutional definitions, at a municipality level. Eventually, referring to those institutional definitions of the affected areas, some alternative methodologies have been suggested to undertake a counterfactual analysis.

Firstly, a cluster analysis has been performed at municipality level, in order to detect major socio-economic differences. Six out of ten typologies comprise both affected and non-affected municipalities. Eventually, a counterfactual analysis based on DID procedure suggests that more manufacturing municipalities have registered the worst performances, in terms of both population and employment losses. Nonetheless, municipalities within cluster #6, #4 and #3 tend to be the most affected ones in comparison to non-affected observations.

Besides punctual observations, these findings are particularly important as they could represent an important basis for more general analyses on the effects produced by an earthquake in the short to medium period. In particular, from policy makers perspective, this analysis returns robust quantitative results to evaluate the loss of economic activity which has been caused by a natural disaster. Thus, it allows the formulation of specific policy interventions.

Furthermore, a similar analysis can be applied to other natural disasters that have occurred to Italy since the WWII. Further studies will apply same methodology that has been developed here to the 1968 Belice earthquake (in Sicily). Thus, a comparison between short and longer effects would deepen general knowledge on natural disaster impact on a sub-regional scale.

References

Abadie, A., Diamond, A., and Hainmueller J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490): 493-505.

Abadie, A., Diamond, A., and Hainmueller, J. (2011). Synth: An R Package for Synthetic Control Methods in Comparative Case Studies. *Journal of Statistical Software* 42(13): 1-17.

Albala-Bertrand, J.M. (1993a). Political Economy of Large Natural Disasters: With Special Reference to Developing Countries. Oxford: Oxford University Press.

Albala-Bertrand, J.M. (1993b). Natural Disaster Situations and Growth: A Macroeconomic Model for Sudden Disaster Impacts. *World Development*, 21(9): 1417–34.

Anderies, J.M., Janssen, M.A. and Ostrom, E. (2004). A framework to analyze the robustness of social-ecological systems from an institutional perspective. *Ecology and Society* 9(1), 18 [online] URL http://www.ecologyandsociety.org/vol9/iss1/art18/.

Austin, P. C., Grootendorst, P., and Anderson, G. M. (2007). A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study. *Statistics in Medicine*, *26*: 734-753.

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

Barone, G. and Mocetti, S. (2014). Natural disasters, growth and institutions: A tale of two earthquakes. *Journal of Urban Economics*, vol. 84: 52-66.

Benni, F. and Brasili, C. (2006). Un indicatore sintetico di attività economica per le regioni italiane, *Rivista di Economia e Statistica del Territorio* 2.

Benson, C. and Clay, E.J. (2004). Understanding the Economic and Financial Impacts of Natural Disasters. 28406. Disaster Risk Management Working Papers 4. Washington, D.C.: The World Bank.

Brusco, S. (1989). *Piccole imprese e distretti industriali*. Torino: Rosemberg & Sellier.

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1): 31-72.

Campbell, D.T., and Stanley, J. (1966). Experimental and Quasi-experimental Designs for Research. Chicago: Rand McNally.

Campbell, D.T., Stanley, J. Caliński, T. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics* 3(1), 1–27.

Cavallo, E. and Noy, I. (2009). The Economics of Natural Disasters: A Survey. Research Department Publications 4649. Inter-American Development Bank, Research Department.

Cavallo, E., Powell, A. and Becerra, O. (2010). Estimating the Direct Economic Damages of the Earthquake in Haiti. *The Economic Journal* 120 (546): 298–312.

Centro di Portici (1981). Situazione, problemi e prospettive dell'area più colpita dal terremoto del 23 novembre 1980. Torino: Einaudi Editore.

Cuaresma, J.C. (2009). Natural Disasters and Human Capital Accumulation. Policy research working paper 4862. Working Paper. Washington, D.C.: World Bank.

De Maria, C. (ed.) (2012). Bologna futuro: il modello emiliano alla sfida del XXI secolo. Bologna: Clueb.

DuPont IV, W. and Noy, I. (2015). What Happened To Kobe? A Reassessment of the Impact of the 1995 Earthquake in Japan. *Economic Development and Cultural Change*, 63(4). 777-812.

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

Istat (2011a). 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.

Kahn, M.E. (2003). The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions. Tufts University.

Kaufman, L. and 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. and Williams W.T. (1966). A Generalized Sorting Strategy for Computer Classifications. Nature 212(5058): 218.

Mosconi, F. (ed.) (2012). Le metamorfosi del "modello emiliano". L'Emilia-Romagna e i distretti industriali che cambiano. Bologna: Il Mulino.

Okuyama, Y. and Chang, S.E. (2004). *Modeling Spatial and Economic Impacts of Disasters*. Berlin: Springer.

Olmos, A. and Govindasamy, P. (2015). Propensity Scores: A Practical Introduction Using R. Journal of MultiDisciplinary Evaluation, 11(25): 68-88.

Pagliacci F., Bertolini P. (2015). Le specificità del sistema agro-alimentare nella ricostruzione post-sisma, DEMB Working Paper Series n. 68. Università di Modena e Reggio Emilia.

Piazzi, V., Pagliacci, F. and Russo, M. (2015). Analisi cluster delle caratteristiche socio-economiche dei comuni dell'Emilia-Romagna: un confronto tra comuni dentro e fuori dal cratere del sisma. DEMB Working Paper Series, 68. Università di Modena e Reggio Emilia.

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/.

Ranuzzini M., Pagliacci F., Russo M. (2015). L'informatizzazione delle procedure per la ricostruzione: prime evidenze dai contributi concessi per le abitazioni, DEMB Working Paper Series n. 71. Università di Modena e Reggio Emilia.

Rosenbaum, P.R. and Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1): 41-55.

Sekhon, J.S. (2011). Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R. *Journal of Statistical Software* 42(7): 1-52.

Strömberg, D. (2007). Natural Disasters, Economic Development, and Humanitarian Aid. *Journal of Economic Perspectives*, 21(3): 199–222.

The United Nations e The World Bank (2010). *Natural Hazards, Unnatural Disasters: The Economics of Effective Prevention*. Washington, D.C.: World Bank Publications.

Tobler, W.R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46: 234–240.

Tol, R.S.J. e Leek. (1999). Economic Analysis of Natural Disaster. In Climate Change and Risk, 308–27. London: Routledge.

UNISDR (2009). Terminology on Disaster Risk Reduction. Geneva: United Nations Office for Disaster Risk Reduction.

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