2012 Emilia earthquake, impact on income inequality

Alegría Burneo 3139736
Cecilia Maria Lanzavecchia 3141585
Maho Yoneda 3126343
Manuela Orsi 3107919

Bocconi University, Milan
December 2022

Abstract

Studies on the effect of natural disasters on inequality are done mostly in developing countries and their conclusions vary hugely, with little applicability to the cases in developed countries. This paper analyzes the effect of the two earthquakes occurred in Emilia Romagna in 2012. In particular, we determine whether the earthquakes had a short-term impact on the region's income inequality. To do so, we utilize difference-in-difference method on measures of income inequality such as Gini coefficient and $90^{th}/10^{th}$ income percentiles ratio, which we generate using administrative data published by INPS. The study finds no evidence that the earthquakes had any significant effect on income equality in the region.

1. Introduction

In May 2012, two major earthquakes struck Emilia Romagna. The first shock hit the area on May 20th with a magnitude of 5.9 on the Richter scale, while the second one occurred on May 29th with a magnitude of 5.8. The events, known in Italy as the 2012 Emilia earthquake, caused 28 deaths, 300 injured, 45 thousand displaced people and damage amounting to 13.2 billion euros[1]. As the name suggests, Emilia Romagna was the epicenter of these earthquakes and, in particular, the provinces of Reggio Emilia, Modena, Bologna and Ferrara were the most affected. Hence, we considered the overall impact of these two earthquakes since they happened very close in time. The area was densely populated, industrialized and it had a high employment rate. 10 thousand municipal contributions have been asked by the affected population, in order to fix damages to more than 27 thousand real estate units. Among those, 20.2 thousand were houses and 7 thousand were destined for commercial and productive activities (shops, offices, craft shops, warehouses).

1.1. Motivation and Research Question

In this paper, we examine the short-term effect of 2012 Emilia earthquake on income inequality in Emilia Romagna. Conclusions from previous studies on the effect of natural disasters vary significantly: one study finds that the Tropical Cyclone Nargis that hit Myanmar in 2008 raised inequality within the affected region (Warr & Aung, 2019), while another concludes that natural disasters that occurred from 1990 and 2013 in Sri Lanka decreased income inequality (Keerthiratne & Tol, 2018). Furthermore, another study finds that 2008 Sichuan earthquake in China did not have any effect on income inequality (Feng, Lu, Nolen, & Wang, 2016). However, it is important to note that the related literature focuses on developing countries, which may have different characteristics specially in terms of infrastructure and institutions, compared to our region of interest. Therefore, we aim at understanding which of the mentioned studies best describes the effect of Emilia Romagna earthquake in the region. We used two different measures for income inequality: the Gini coefficient[2] and the 90th/10th income percentiles ratio[3]. Our hypothesis is that the earthquake would increase income inequality. This would happen because we believe that middle-high income people are usually in high skilled positions, which are not easily substitutable and could also be performed remotely, thus they are less likely to lose their job as an effect of the earthquake. On the contrary, low-income workers tend to have more physical jobs which cannot be done in distance (e.g., waiters, cashiers, warehouse workers), which means that in the short run they are more likely to lose their job if their working place gets destroyed. Therefore, middle-high income workers would continue to earn the same income, while low-income workers may earn less leading to an increase in income inequality.

[1] A sei anni dal sisma. (n.d.). Regione Emilia-Romagna

[2] The Gini coefficient is the most popular statistical measure of economic inequality which measures the dispersion of income within a population. The Gini coefficient is a value between 0% and 100% with 0% representing a perfectly equal society with all income equally shared and 100% reflecting maximum inequality.

[3] The 90th/10th income percentiles ratio compares the income earned by individuals at the 90th percentile to the earnings of workers at the 10th percentile.

2. Methodology and Data

2.1. Data

We started our analysis on administrative data provided by INPS. We organized the dataset as panel data, consisting of observations of 165,837 individuals living in all 20 Italian regions over a period of time from 2005 to 2016. In our research, we decided to narrow the time frame considering the effect of the earthquake until 2014 to focus on the short-term effects and to avoid a possible attenuation effect due to a long-term prospective. Among the variables in the dataset, the following ones at the individual level were considered: income, year, employment, qualification, gender, region code[4] and ateco 2007 (which is a variable for the classification of economic activities by industry). The income variable refers to annual labor income, which we used to create the Gini coefficient and the $90^{th}/10^{th}$ income percentiles ratio, both at the regional level by year, which will grasp income inequality. Finally, we created dummy variables for being female, years, qualifications, economic sectors (primary, secondary, and tertiary), and macro areas (North, Center, and South Italy).

2.2. Empirical Model

In order to understand whether the earthquake had a significant impact on income inequality in Emilia Romagna, we used the difference in difference method. This is because it allows to isolate the effect of the shock on income inequality in Emilia Romagna, which we can then use to compare with other Italian regions. In our research the "treatment" corresponds to the 2012 earthquake suffered by Emilia Romagna. As "before treatment" period we considered 2005-2011 and as "after treatment" 2012-2014, in order to focus on the short-term effect. Therefore, the treated group is Emilia Romagna, and the control group is composed by all the other Italian regions.

In the first place we must assess whether the assumptions of the model are satisfied:

- Exogeneity
- Absence of spillover effects of the treatment
- Parallel trends in terms of Gini coefficient and $90^{th}/10^{th}$ income percentiles ratio of the two groups in absence of the treatment.

Even though Emilia Romagna has experienced other earthquakes in the past, the area's seismic risk is not high[5]. Therefore, since the area is not frequently affected by earthquakes, we can consider the earthquake as an exogenous event, thus there are no problems arising from endogeneity.

The absence of spillover effect, also known as contamination effect, refers to the assumption that the treatment affects only those who are treated. In our case, this assumption may not be satisfied due to two main effects on regions close to Emilia Romagna: the direct effect of damages from the earthquake, and the indirect effect of migration. For the former effect, the earthquake could have hit nearby regions too, however, in this case this is not a problem since the damages suffered by them are not significant[6]. For the second effect, income distribution of neighboring regions may have been affected by the

[4] COD_REG
[5] Seismic risk in Emilia-Romagna. (n.d.). Environment
[6] I danni provocati. (n.d.). Regione Emilia-Romagna

earthquake due to immigration of people who have lost their houses or jobs from Emilia Romagna. Therefore, to try to control for the effect of migration, we followed a robustness check by creating another regression excluding the observations for neighboring regions, namely: Lombardia, Veneto, and Toscana.

Lastly, the parallel trends assumption refers to the idea that the income inequality measures considered for both the treatment and control group would follow the same trend in the absence of the earthquake over time. To check this assumption, we built two graphs comparing respectively the trends of the Gini coefficient and the $90^{th}/10^{th}$ income percentiles ratio of Emilia Romagna with the average trend of all other regions. From
Figure 1 one can see that the pre-treatment trends of both inequality measures are sufficiently parallel, with a slight divergence in the post-treatment period. Thus, we decided to extend our research by limiting the control group to regions more similar to Emilia Romagna, specifically by excluding southern regions.

Moving to the difference in difference regressions, our two models have as dependent variable the Gini coefficient and the $90^{th}/10^{th}$ income percentiles ratio respectively, both measured in region i and period t:

1) $Gini_{it} = \beta_0 + \beta_1 After_t + \beta_2 (After_t * ER_i) + \beta_3 reg_i + + \beta_4 dq^*_{it} + \beta_5 ds^*_{it} + \beta_6 fem_{it} + \varepsilon_{it}$

2) $Ratio90/10_{it} = \beta_0 + \beta_1 After_t + \beta_2 (After_t * ER_i) + \beta_3 reg_i + \beta_4 dq^*_{it} + \beta_5 ds^*_{it} + \beta_6 fem_{it} + \varepsilon_{it}$

- $After_t$ is a dummy variable set to 1 if the observation is after the earthquake.
- $After_t * ER_i$ is our interactive variable set to 1 for the observations of Emilia Romagna and after the earthquake, thus β_2 captures the difference in difference effect.

As control variables for the difference in difference regression we included:
- reg_i that is the fixed effect for each region;
- dq^*_{it} is the percentage of people in executive positions, office workers, manual workers and managers in region i and period t;
- ds^*_{it}: percentage of people employed in the primary sector / secondary sector (tertiary sector is not included to avoid multicollinearity), in region i and period t;
- fem_{it} is the observed percentage of women in region i and period t is the percentage of women in region i and period t.

The employment variable was excluded from our regressions as control variable for two possible reasons: it is likely to be correlated with the shock and it could be an *in-period variable*, denoting that it would be measured in the same period as the Gini coefficient or the $90^{th}/10^{th}$ income percentiles ratio.

Furthermore, to make our analysis more precise we decided to implement a synthetic control method to create a better counterfactual to see whether the evolution of the considered inequality measures in Emilia Romagna would have been different in the absence of the earthquake, while relaxing the assumption of parallel trends. To construct the synthetic control, we firstly used as predictor variables the Gini coefficient together with the percentage of employed people by region and year, the percentage of people with different qualifications, the percentage of workers in the three sectors, percentage of female and the macro areas, all from 2005 to 2011. Secondly, we constructed another synthetic control with the $90^{th}/10^{th}$ income percentiles ratio, instead of the Gini coefficient, and all the predictors mentioned before. All these variables are used to model as closely as possible the economic structure of Emilia Romagna and simulate the distribution of income by assigning weights to similar regions. For instance, we believe that the percentage of people in managerial positions can be used as a proxy for high income households, which is a possible driver of inequality. Moreover, we used the percentage of workers in each of the sectors (primary, secondary, and tertiary), since on average agriculture workers have lower wages than people working on services, thus being an important factor for income distribution in the region. Finally, we considered the dummy variables for macro areas as on average the regions in each of these areas have similar geographical and socio-economic characteristics, and the macro areas differ between them as seen in Figure 2.

3. Results and robustness checks

The results of our five regressions of the Gini Coefficient and the $90^{th}/10^{th}$ income percentiles ratio are summarized respectively in Table 1 and Table 2. The control group of each regression is described under the results tables. We found that in both cases our coefficient of interest, the one of the interactive variable (ER_after), was not significant in any of the five regressions. Among the five regressions, the first regression has the lower R^2 in case of both inequality measures, and this is because it does not control for regional characteristics. From the second regression onwards, we controlled for region characteristics such as qualification, employment in primary and secondary sectors, female percentages and regional fixed effect (absorbed, see Appendix). In the second regression and in case of both inequality measures, some controls were significant (e.g., employment in secondary sector and female percentages). This means that they have an effect on the Gini coefficient / the $90^{th}/10^{th}$ income percentiles ratio[7]. For example, an increase in the employment rate in the secondary sector by one percentage point leads to a decrease in the Gini coefficient of 0.00328, and a decrease in the $90^{th}/10^{th}$ income percentiles ratio of 0.541[8].

However, it is possible that the control group we are using is biased since it considers all Italians regions which are very different in nature. Therefore, we decided to narrow our control group by excluding southern regions. Thus, in the third regression the control group is composed by northern and central regions, since Emilia Romagna's inequality measures behave more similarly to these macro areas rather than South Italy, as one can see in Figure 2. Also in this case our coefficient of interest, the interactive variable, is not significant. Moreover, in this model specification, the percentages of people employed in the primary and secondary sectors are significant in case of both inequality measures. Subsequently, we decided to run a synthetic control to deepen our research and find a better control group. By running a synthetic control with all regions, including the Gini coefficient leads to have five regions relevant for our purpose. In order of importance: Veneto (0.383), Lombardia (0.308), Valle d'Aosta (0.167), Piemonte (0.085) and Abruzzo (0.058). On the other hand, including the $90^{th}/10^{th}$ income percentiles ratio leads to the following weights: Veneto (0.402), Lombardia (0.247), Piemonte (0.125), Friuli Venezia Giulia (0.107), Valle d'Aosta (0.083)[9]. The regions that were given the higher weights were Veneto and Lombardia, thus the inequality measures considered led to consistent and similar results. Our fourth regression, yielded non-significant results regarding the coefficient of interest. Regarding the control variables used in this regression, only the percentage of employment in secondary sector is significant for both inequality measures. Finally, to control for the effect of migration, we ran another synthetic control excluding the nearby regions of Lombardia, Veneto, and Toscana, as they share the longest borders with Emilia Romagna. Now, more weights are attached to Valle D'Aosta (0.775) followed by Piemonte (0.14), Trentino Alto Adige (0.032), Friuli Venezia Giulia (0.025) and Marche (0.028), in case of the Gini coefficient. Instead, in case of the $90^{th}/10^{th}$ income percentiles ratio weights were given to Piemonte (0.662), Friuli Venezia Giulia (0.123), Molise (0.074), Umbria (0.072), Marche (0.07). By running our fifth regression, we found that again only the percentage of employment in the secondary sector is significant across both inequality measures and a non-significant coefficient for the interactive variable.

Although the effect of the earthquake is not significantly different from zero in any of our regressions, the R^2 is high in all regressions controlling for regions characteristics. In case of the Gini coefficient, the variables included are able to explain from 84.4% to 93.2% of the variability of the dependent variable. Instead, in case of the $90^{th}/10^{th}$ income percentiles ratio this range is from 83.5% to 90.2%. This high degree of explained variability of the dependent variables means that the choice of the independent variables was reasonable and quite exhaustive.

[7] All interpretations of the coefficients from our results consider the marginal change in the variable of interest keeping all other variables constant.

[8] Both changes are with respect to the measures for the tertiary sector, since it is omitted.

[9] Liguria (0.014), Marche (0.002).

4. Conclusion

Our paper shows that the 2012 Emilia-Romagna earthquake's impact on inequality was non-significant denoting that inequality in Emilia-Romagna has not significantly changed because of the 2012 earthquake compared to inequality in other Italian regions.

Moreover, the results are not in line with our initial hypothesis that the earthquake would have led to an increase in inequality. The null effect of the earthquake could be explained by two possible factors: state interventions and attenuation effect. For the former, some Italian articles, already cited earlier in this paper, show that there have been some municipal contributions that account for €6.4 billion for reconstruction of houses and production activities which were destroyed by the earthquake. Moreover, Cassa Integrazione, which is a state aid used to support those employees that, in our research, lost their jobs due to the earthquake, was granted for an amount of almost €70 million to support employed and self-employed people (DL74/L122). Therefore, it can be the case that the policy intervention in Emilia Romagna was effective enough to counteract the earthquake effect. Moreover, as far as the attenuation effect is concerned, only few cities were strongly affected by the earthquake while the majority of cities within Emilia-Romagna did not experience strong destructions at all. However, by having data just at a regional level we could not detect the different impact of the earthquake within the region. Therefore, it is possible that data from heavily affected cities were attenuated by the non-affected ones. Furthermore, looking at Figure 6 we notice that the Gini coefficient followed a stable trend after 2012, with very little variations that could have been due to the natural evolution of the region over time. On the other hand, the $90^{th}/10^{th}$ income percentiles ratio had a more dynamic trend across the years considered: it had an upward trend until 2012 and then it diminishes. However, this paper, within its limitations, concludes that the earthquake does not seem to be the reason behind the drop after 2012.

4.1 Limitations

Our paper has some limitations due to the nature of the dataset. Firstly, we did not have disaggregated data for income. Specifically, we had no monthly data and no province/city data. Thus, with only annual income data available we were not able to identify the immediate effect of the earthquake on income inequality in following months. Furthermore, having had data at the province/city level would have allowed us to control for the attenuation effect, meaning for differences in damages within the region and inter-regional migration according to which we could better control for the spillover effect. Moreover, we had no data for migration due to the earthquakes that would have allowed us to control better for the spillover effect driven by migration. In fact, in our studies, we tried to control for the spillover effect by running a robustness check, but we cannot be sure that migration actually happened or in case it did happen, if people moved to the regions we excluded or others. We were not even able to detect any internal migration within Emilia-Romagna region. Finally, we had to construct the Gini coefficient by using income and not capital data that were not available. Therefore, the Gini coefficient we constructed grasps income inequality rather than wealth inequality. Since an earthquake usually has a higher effect on physical capital, leading people to be more likely to lose capital rather than income (from their job), our Gini coefficient index does not grasp the physical damages. On the other side, some limitations derive from the nature of the $90^{th}/10^{th}$ income percentiles ratio: it ignores information about incomes in the middle of the income distribution and doesn't use information about the distribution of income within the top and bottom percentiles. Therefore, other indexes could have been used to detect stronger impacts of the earthquake in terms of wealth inequality.

4.2 Further Research

To develop our research further, it could be relevant to conduct the same research by adding the missing data previously stated: income data at monthly and city level as well as migration data. Furthermore, as long as capital data are available, it may be that more significant results can be found by considering capital instead of income, thus studying for wealth inequality rather than income inequality after the struck of an earthquake.

Additional research could be made by comparing the impact of an earthquake, among regions or countries with different institutions to see whether a country with poor institution may react differently in terms of policy intervention. In fact, one study finds that countries with better institutions experience less victims and lower economic losses from natural disasters (Raschky, 2008). Besides, we think that conducting our research on countries in which there is already a high initial inequality and in which an earthquake could have very catastrophic effect on low-income people could lead to significant results. As a matter of fact, in developing countries there is a high difference in the infrastructure in which low-income and high-income people live. Therefore, the struck of an earthquake could have an extremely high effect on inequality.

5. Appendix

5.1. Description of construction of dummy variables

In particular, to generate the dummy variables for the primary, secondary and tertiary sectors we used the ateco 2007 variable, which takes value ranging from 1 to 99, with each number representing an economic sector. We grouped under primary sector the ateco from 1 to 8, being agriculture, mining and extraction sectors; under secondary sector we grouped the ateco from 9 to 35, being manufacturing and energy providing sectors; finally, under tertiary sector we grouped ateco from 36 to 99, being the services sectors[10]. In order to avoid multicollinearity, for each set of dummy variables we dropped one of them: other qualification ("altro"), year 2012, tertiary sector and South Italy.

Regarding the methodology, since the dataset contained many groups, when running our regressions, we used the command "areg" absorbing the categorical variable region code, which was thus included in the regressions as if it was specified by dummy variable. This is why in the results tables regional fixed effects do not appear.

5.2. Description of construction of graphs

We constructed various graphs (Figure 1 to 6) which represent the evolution of each of the coefficients (Gini coefficient and 90/10th income percentiles ratio) throughout the years. In particular we have years on the x-axis, from 2004 to 2014, and the values of the indices on the y-axis. To visualize the effect of the earthquake we added a line in 2011 representing the threshold for the post-treatment period. We chose to position it in 2011 due to the fact that the earthquake hit the region in May 2012 and given that we only have annual data for income, we would observe the immediate effects of the earthquake on the reported data for income of 2012.

[10] Ateco 2007 a 2 cifre. (n.d.).

5.3. Tables

Table 1

	(1) gini	(2) gini	(3) gini	(4) gini	(5) gini
R squared	0.0445	0.932	0.958	0.900	0.844
after 2012	0.0123*	0.00741**	-0.000462	-0.00155	0.00155
	(0.00482)	(0.00231)	(0.00192)	(0.00244)	(0.00280)
Emilia Romagna after 2012	-0.00683	0.00148	0.00364	0.00481	0.00413
	(0.00587)	(0.00333)	(0.00296)	(0.00294)	(0.00388)
apprentice (%)		0.0244*	0.00413	0.00874	0.0139
		(0.0113)	(0.00982)	(0.0117)	(0.00843)
executive (%)		0.0400**	0.0198	0.00534	0.0162
		(0.0134)	(0.0126)	(0.0179)	(0.0131)
office worker (%)		0.0233*	0.00481	0.0140	0.0201*
		(0.0113)	(0.00981)	(0.0125)	(0.00924)
manual worker (%)		0.0252*	0.00683	0.0139	0.0205*
		(0.0111)	(0.00946)	(0.0119)	(0.00825)
manager (%)		0.00692	-0.00642	-0.00419	-0.00308
		(0.0123)	(0.0115)	(0.0122)	(0.00655)
employment in primary sector (%)		0.0106	0.0217***	0.0178**	0.00664
		(0.00584)	(0.00489)	(0.00582)	(0.00794)
employment in secondary sector (%)		-0.00328***	-0.00363***	-0.00351***	-0.00240***
		(0.000469)	(0.000485)	(0.000479)	(0.000551)
female(%)		0.00405***	0.000869	-0.00184	0.00263
		(0.00108)	(0.00101)	(0.00105)	(0.00137)
Observations	200	200	120	60	60

Standard errors in parentheses

(1) regression without controls, control group is all other Italian regions;

(2) regression with controls, control group is all other Italian regions;

(3) regression with controls, control group is North and Center regions of Italy;

(4) regression with controls, control group is synthetic Emilia Romagna;

(5) regression with controls, control group is robust synthetic Emilia Romagna;

(6) regression with controls, control group is robust synthetic Emilia Romagna excluding Valle d'Aosta;

All regressions except for (1) are controlled for baseline gini (2011);

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2

	(1) ratio 90/10	(2) ratio 90/10	(3) ratio 90/10	(4) ratio 90/10	(5) ratio 90/10
R squared	0.518	0.835	0.876	0.864	0.902
after 2012	2.250***	0.728*	0.452	-0.0238	-0.101
	(0.337)	(0.342)	(0.385)	(0.435)	(0.371)
Emilia Romagna after 2012	-0.479	0.756	0.925	0.745	0.452
	(0.889)	(0.530)	(0.492)	(0.568)	(0.549)
apprentices (%)		3.435*	3.622	3.995	-2.206
		(1.590)	(1.850)	(2.014)	(1.144)
executives (%)		3.540	3.599	3.740	-2.759
		(2.034)	(2.227)	(3.380)	(1.975)
office workers (%)		3.721*	4.241*	5.367*	-1.206
		(1.578)	(1.807)	(2.074)	(1.209)
manual workers(%)		4.260**	4.845**	5.640**	-0.335
		(1.544)	(1.738)	(1.954)	(1.128)
managers (%)		2.094	2.112	3.005	-1.822
		(1.723)	(1.933)	(2.128)	(1.307)
employment in primary sector (%)		2.472***	2.220**	1.928	0.581
		(0.694)	(0.691)	(0.991)	(1.447)
employment in secondary sector (%)		-0.541***	-0.547***	-0.566***	-0.298***
		(0.0597)	(0.0574)	(0.0542)	(0.0699)
female (%)		0.487**	0.429**	0.503	0.657**
		(0.166)	(0.158)	(0.258)	(0.235)
Observations	200	200	120	60	60

Standard errors in parentheses

(1) regression without controls, control group is all other Italian regions;

(2) regression with controls, control group is all other Italian regions;

(3) regression with controls, control group is North and Center regions of Italy;

(4) regression with controls, control group is synthetic Emilia Romagna;

(5) regression with controls, control group is robust synthetic Emilia Romagna;

All regressions except for (1) are controlled for baseline gini (2011);

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.4. Figures

Figure 1

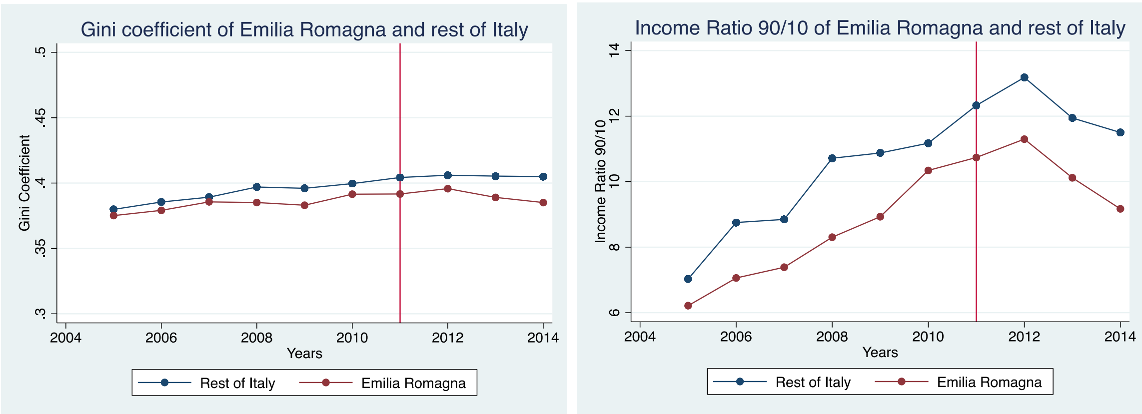


Figure 2

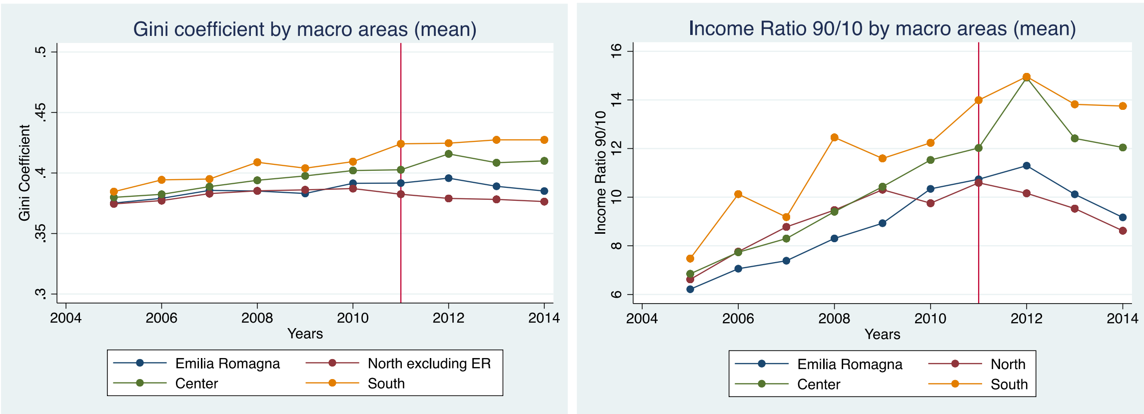


Figure 3

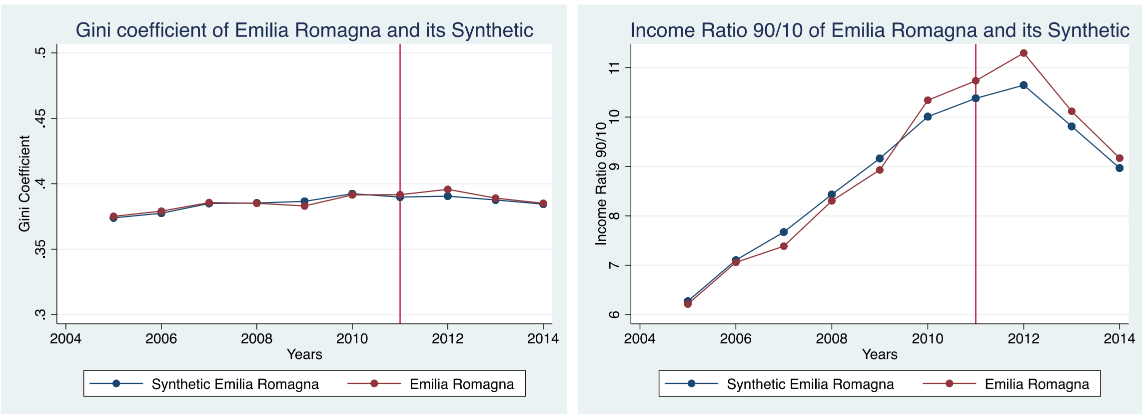


Figure 4

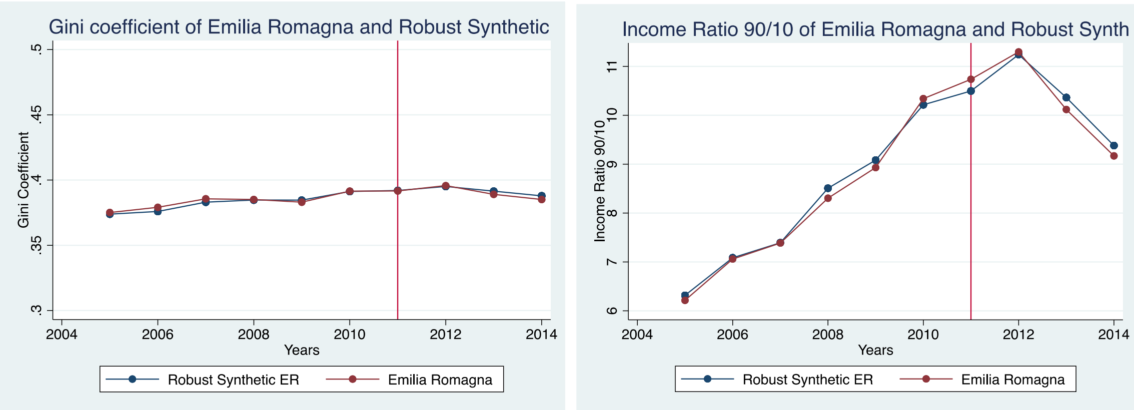


Figure 5

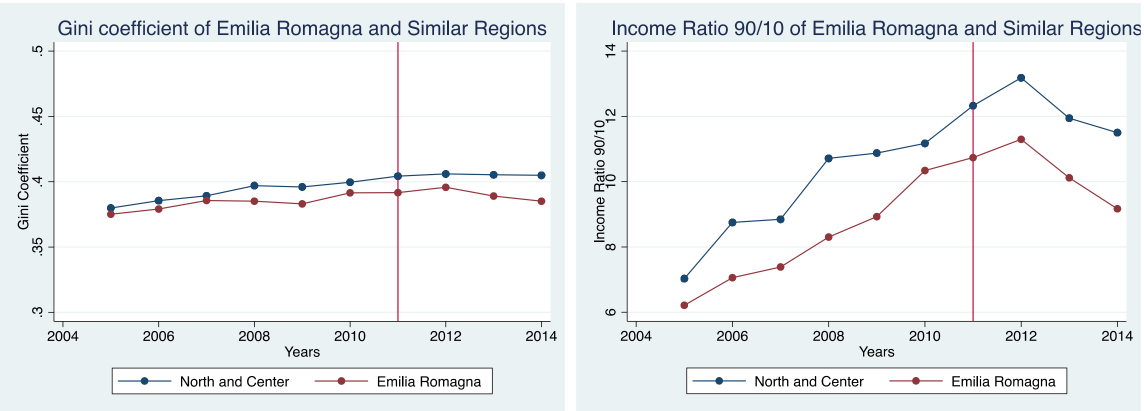
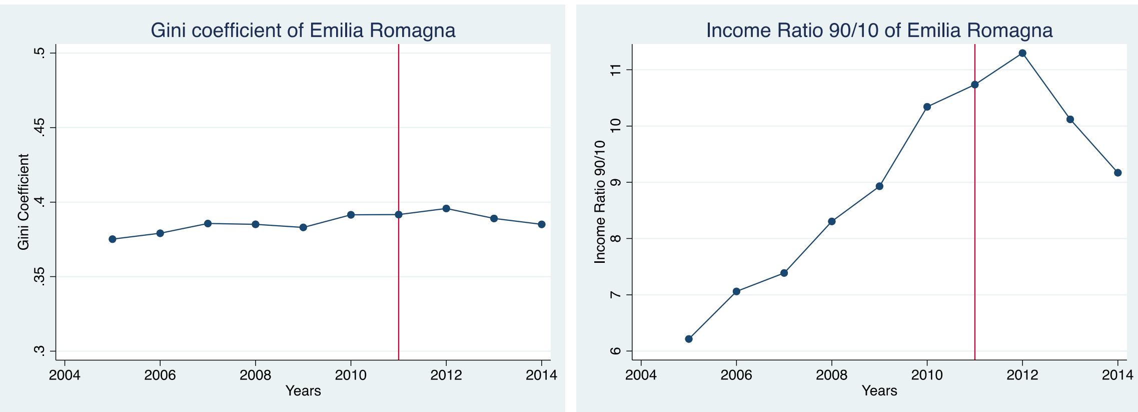


Figure 6



Reference

Warr, P., & Aung, L. L. (2019). Poverty and inequality impact of a natural disaster: Myammar's 2008 cyclone Nargis. World Development, 122, 446-461.

Keerthiratne, S., & Tol, R. (2018). Impact of natural disasters on income inequality in Sri Lanka. World Development, 105, 217-230.

Feng, S., Lu, J., Nolen, P., & Wang, L. (2016). The effect of the Wenchuan earthquake and government aid on rural households. In IFPRI (pp. 11-34).

Raschky, P. A. (2008). Institutions and the losses from natural disasters. Natural hazards and earth system sciences, 8(4), 627-634.