

Impacts of large natural disasters on regional income

Jianhong E. Mu¹ · Yong Chen¹

Received: 7 October 2015 / Accepted: 19 May 2016 / Published online: 31 May 2016
© Springer Science+Business Media Dordrecht 2016

Abstract In this study, we analyze impacts of large natural disasters on county per capita income between 1990 and 2012 in the Contiguous United States. Using a difference-in-differences model with fixed effects and controlling for serial correlation, we find that the incidence of large disasters in a county significantly reduces its income as compared to its neighboring counties.

Keywords Natural disasters · Regional income · Difference-in-differences

JEL classification R1

1 Introduction

Economic activities are frequently interrupted by natural disasters, like the 2005 Hurricane Katrina and severe winter storms in recent years. While most existing economic studies examine the impact of disasters on the national economy (see Cavallo and Noy (2011) and Kousky (2014) for a review), more and more researches start looking at regional economic impacts. In this study, we focus on the impact on per capita income. In this regard, most existing studies find that natural disasters have a significant negative impact in the

Electronic supplementary material The online version of this article (doi:10.1007/s11069-016-2372-3) contains supplementary material, which is available to authorized users.

✉ Jianhong E. Mu
mujh1024@gmail.com

Yong Chen
yong.chen@oregonstate.edu

¹ Department of Applied Economics, Oregon State University, Corvallis, OR 97331, USA

short-run (Strobl 2011; Xiao 2011; Coffman and Noy 2012; Anttila-Hughes and Hsiang 2013).¹ However, the long-run impact on regional income is rarely discussed except in Coffman and Noy (2012) and Xiao (2011).

Xiao (2011) examines impacts of the 1993 Midwest flood on local economies and finds that per capita income drops significantly in the year of the event. In the long run, the impact on per capita income is negative but not significant.² Coffman and Noy (2012) apply the synthetic control method to examine the economic impact of hurricane Iniki on the Hawaiian island of Kauai. They find that hurricane Iniki has mainly affected population. It significantly reduces per capita income in the short run but not in long run. The significant short-run impacts and the insignificant long-run impacts seem to suggest that policy makers should focus on the short-run instead of the long-run.

Most existing regional studies on disaster impacts focus on one particular type of natural disasters, such as hurricanes (Anttila-Hughes and Hsiang 2013; Belasen and Dai 2014; Coffman and Noy 2012; Deryugina 2013; Strobl 2011), flooding (Husby et al. 2014; Xiao 2011) and wildfires (Nielsen-Pincus et al. 2014). The exception is Vu and Hammes (2010). They investigate impacts of multiple natural disasters on annual output and output growth over the period 1995–2007 in China. One concern about focusing on one particular type of disasters is the neglect of disaster incidences of other types, which raises a question about the validity of the control group. Should counties stricken by tornados but not by flooding be included in the control group in a study of flooding?

To address this concern, we investigate the disaster impact on per capita income considering multiple types of disasters occurred in the Contiguous US from 1990 to 2012. After controlling for the potentially long-lasting impact from previous disasters, we construct counterfactual control groups for counties hit by large natural disasters. We collapse the time series data into pre- and post-event periods to produce consistent standard errors (Bertrand et al. 2004). After controlling for the potential serial correlation problem in the difference-in-differences model, we find that comparing with the neighboring counties, the per capita income in counties directly stricken by large natural disasters are significantly reduced in the long-run.

The remainder of the paper is organized as follows. Section 2 describes the econometric model. Section 3 describes the data and the assignment of the treatment and control groups. Section 4 presents results and discusses robustness checks and Sect. 5 concludes the paper.

2 Econometric models

It is necessary to clarify notations and terminologies used in this paper in order to facilitate the discussion. When a large natural disaster³ occurs, we say a treatment event has occurred. Because the dataset includes disasters occurred in multiple years, it is important to differentiate two related but different concepts: year and period. In this study, the year

¹ Instead of using income, Belasen and Polachek (2008) and Deryugina (2013) use earnings and find different results. Balasen and Polachek (2008) find that hurricanes have a positive impact on the growth rate of earning per worker in the short run. Deryugina (2013) finds no significant impact on the level of average earnings in the short run but a positive impact in the long run.

² Xiao (2011) investigates the long-run impacts of flooding for both low and high damages using both one and two controls to match each treatment. The long run impact is only significant for the flooding with low damage and using one-treatment-to-two-control match.

³ The definition of a large disaster will be discussed in the next section.

that the treatment event occurs is called the treatment year (*yr*). The year preceding the treatment year ($yr - 1$) is referred to as the pre-event period. In order to investigate whether a disaster affects county per capita income j years after the treatment event, we define j as the impact duration, and call the j th year after the treatment year as the post-event period. For any given disaster, the post-event period refers to a different calendar year if the impact duration is different. For example, if a disaster occurred in 1997, the treatment year is 1997. If the impact duration of interest is 10 years ($j = 10$), the corresponding post-event period refers to the calendar year 2007. According to Cavallo et al. (2013), any impact observed with impact duration greater (less) than 5 years is considered as the long (short)-run impact. To avoid confusion, in the rest of the paper, we use “year” only for the calendar year and “period” only for the pre- or post-event period.

Because the serial correlation in the economic outcomes may underestimate the standard deviation of the estimators, and overestimate significance levels, we collapse the times series data into the pre- and post-event periods as suggested by Bertrand et al. (2004). For a given impact duration j , let y_{it} be the real per capita income for county i in period t . The regression model is:

$$\ln y_{it} = \beta_1 DID_{it} + \beta_2 X_{it} + \beta_3 Period_t + c_i + u_{it} \quad \text{for } t = 0 \text{ or } 1 \quad (1)$$

where DID_{it} is a dummy variable that equals one for counties in the treatment group in the post-event period, and zero otherwise. X_{it} is a vector of weather variables including annual precipitation and mean temperature, as well as their squared terms to control for nonlinear effects. c_i is the county fixed effect used to control for unobserved county characteristics.⁴ $Period_t$ is the time fixed effects to capture common changes from the pre-event period to the post-event period. u_{it} is the idiosyncratic error. β_1 estimates the mean income difference between the constructed treatment and control groups due to the incidences of large disasters.

3 Data

3.1 Data description

The data on natural disasters come from the Spatial Hazard Events and Losses Database for the US (SHELDUSTM Version 12.0), which is maintained by the Hazards and Vulnerability Research Institute (HVRI) at the University of South Carolina.⁵ SHELDUS is a county-level dataset covering 18 types of natural disasters including avalanche, coastal, drought, earthquake, flooding, fog, hail, heat, hurricane/tropical storm, landslide, lightning, severe storm/thunder storm, tornado, tsunami/seiche, volcano, wildfire, wind, and winter weather. This dataset covers natural disasters occurred from January 1960 to present.

⁴ If a county shows up in the treatment group twice, two county fixed effect coefficients are specified. For instance, if a county is hit by a disaster in 1994 and another one in 2005, two fixed effect coefficients are specified for that county to capture the potential changes occurred in the county between 1994 and 2005.

⁵ The main data source of this dataset is the “Storm Data and Unusual Weather Phenomena” from the National Climatic Data Center (NCDC). Due to NCDC’s change in reporting procedures, every event listed in NCDC’s storm dataset that had exact damage values assigned was entered into the database from 1995 onward. However, only selected events with property or crop damage higher than \$50,000 are recorded from 1990 to 1995 (HVRI 2013).

In this research, the direct economic damage is measured by the value of property damage. However, property damages are not directly comparable across regions due to the regional differences in the economic development and living costs. For instance, the property damage from the destruction of a small house in Manhattan, New York, can be very different from the destruction of the same house in a small rural county. To address these differences across regions, we use the concept of relative property damage,⁶ which is defined as the property damage divided by the per capita income in the pre-event period. It is used to measure the property damage as a multiple of average county income. For example, a relative damage value of ten means that the property damage is equivalent to ten times the county per capita income in the pre-event period. The use of relative damage is similar to the normalization using country GDP in Cavallo et al. (2013).⁷ It also allows us to compare direct damages from disasters of different types.⁸

Figure 1 shows the mean and moving average of the relative damage from 1969 to 2012. It is evident that disasters with large relative damages happen more frequently after 1990, suggesting a fatter upper tail in the distribution of relative damage. This is similar to the observation in Misiewicz (2011). Because of this structural change, we only consider disasters that occurred after 1990. The spatial distribution of relative damages in 2005 is plotted in Fig. 2. Most counties with the highest relative damage are located in Louisiana, Mississippi, Alabama and Florida where Hurricane Katrina occurred.

Annual data on county per capita income are available for the period between 1969 and 2012 from the US Bureau of Economic Analysis (BEA).⁹ Nominal values of per capita income are converted to real values using the US consumer price index (1982–1984 = 1) from the Bureau of Labor Statistics. Because real per capita income exhibits time trend and is non-stationary, we use the log real per capita income in our analyses.¹⁰ From here on, we will refer the log real per capita income simply as income.

Annual weather variables are included because evidences show that temperature and precipitation can significantly affect regional economic outcomes (Brown et al. 2013; Dell et al. 2012; Horowitz 2009). The station-wide daily mean temperature (in centimeters) and precipitation (in Celsius degrees) are obtained from the Global Historical Climatology Network (GHCN) database at the National Oceanic and Atmospheric Administration (NOAA). We then aggregate daily weather data to get county-level annual total precipitation and mean temperature.

County rural–urban continuum codes are used in this analysis. They are updated decennially by the Economic Research Service (ERS) of the U.S. Department of

⁶ We do not use crop damage because damage amounts are mainly based on self-report and are potentially influenced by government subsidy and crop/livestock insurance programs, whereas the property damages are based on insurance reports.

⁷ We try to normalize the property damage by county GDP of the previous year as in Cavallo et al. (2013). We find systematic bias in the sample selection, because it tends to select small rural counties into the treatment.

⁸ While physical measurements of disaster intensity like wind speed for hurricanes are preferred, these measures are not generalizable in the sense that they may not be applicable to other disaster types or regions. As a result, we use the concept of relative damage to measure the direct damages from multiple disasters.

⁹ BEA also provides annual economic data like income transfers, employment and population. However, all these variables are endogenous to income. The population data are more problematic because it is essentially imputed from the decennial census data, which tend to smooth out short-run impacts. Consequently, none of these variables are included in the baseline analysis.

¹⁰ The inverse Chi-squared statistic is 6453.6 (p value = 0.001), which rejects the Fisher-type unit-root test, indicating that at least one panel is stationary.

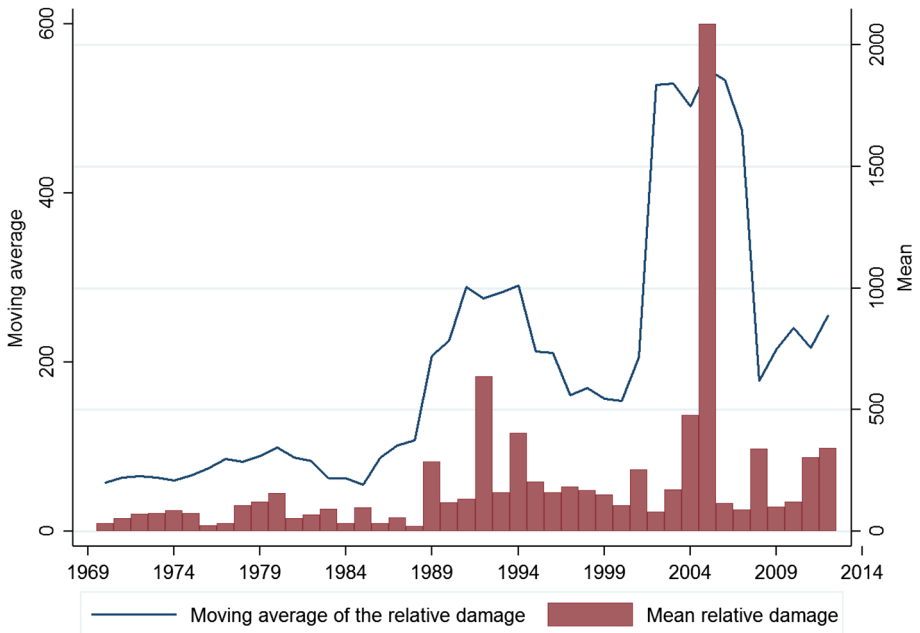


Fig. 1 Temporal distribution of disaster relative damages from 1969 to 2012

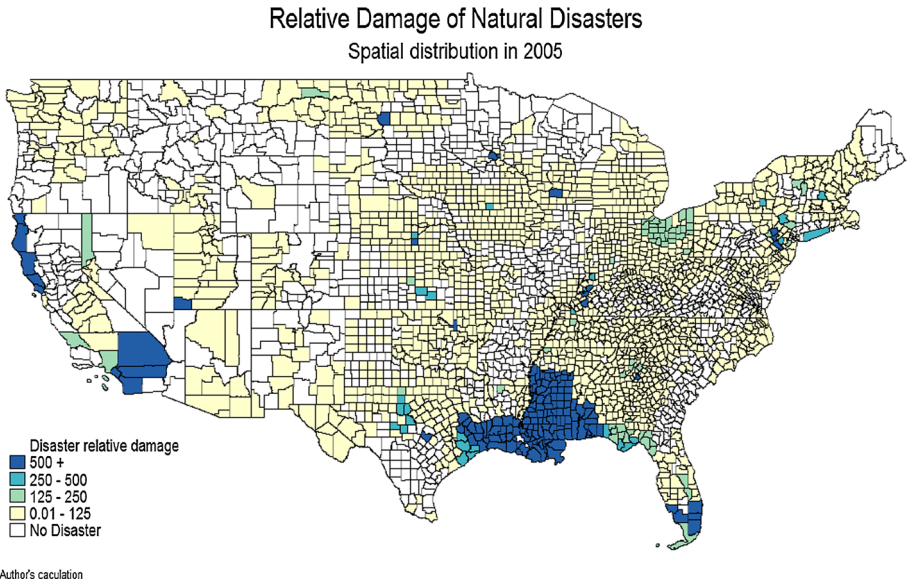


Fig. 2 Spatial distribution of disaster relative damage in 2005

Agriculture in 1983, 1993 and 2003. Following the definition of ERS, a county with rural–urban continuum codes less than or equal to three is considered as a metropolitan county. Otherwise, it is a non-metropolitan county.

3.2 Assignment of the treatment and control groups

We define large natural disasters as disasters with relative damage exceeding some cutoff values, and denote the cutoff value as k^{11} , which is chosen based on the probability distribution of relative damages from all disasters occurred between 1990 and 2012 in the US. For example, $k = 250$ means the disaster damage exceeds 250 times the per capita income of the previous year, which approximately corresponds to the 95th percentile of the relative damage distribution. If a county is hit by a natural disaster generating a relative damage greater than or equal to this cutoff value, the county is called a candidate treatment county with cutoff value $k = 250$.

The quality of a difference-in-differences estimation hinges on the quality of the control group. In this study, due to the lack of exogenous annual county-level data, the quality of the difference-in-differences estimation in essence depends on the degree to which the selected control group can help to control other confounding factors. We use county fixed effects to control the time-invariant unobserved county-specific characteristics. In order to control the unobserved time-varying factors, we resort to the Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970) and construct the control group using neighboring counties.

We define the original dataset as the raw data. A county that receives a treatment event is called a candidate treatment county. A county sharing a common border with the candidate treatment county is called a candidate control county. In order to qualify as a control county, the candidate control county should satisfy the following requirements. Firstly, no treatment event occurs in the county in the same treatment year. This implies that the control county either has no disaster in the treatment year or when a disaster occurs, its relative damage is smaller than the cutoff level. Secondly, the county should have the same metropolitan and non-metropolitan classification as the candidate treatment county according to the rural-urban continuum code. Because metropolitan and non-metropolitan counties can differ significantly in the size and economic structure, this requirement avoids using a metropolitan county as a control for a non-metropolitan treatment county or vice versa. Thirdly, no treatment event should occur in a candidate treatment or control county in the 15 years before the treatment year, which ensures that the pre-event county income of the treatment or control group is not contaminated by previous treatment events. This implicitly assumes that the impact of large natural disaster does not last for more than 15 years. The sensitivity of the estimation results to this assumption will be discussed later. Finally, when we investigate whether the impact lasts for j years (i.e., impact duration equals to j), we also require that a candidate treatment or control county has no treatment event in the j years following the treatment year. This ensures that the county income in the post-event period is not contaminated by treatment events happened after the treatment year. This also implies that the sample size decreases when we try to capture long-run effects. For instance, if the cutoff value equals 150, separate datasets are created for each value of the impact durations ($j = 1, 2, \dots, 10$). Other things equal, the sample size for $j = 1$ is larger than that for $j = 10$, because the latter requires that both the treatment and control group should be free of treatment events for more years.

¹¹ We do not use the sum of relative damages from multiple disasters occurred in the same year, because it may blur the distinction between the treatment and control groups. If the sum is used, it is possible that the treatment and the corresponding control counties may be hit by some common disasters due to the spatial average problem discussed later. Moreover, SHELUS does not record natural disasters with property damage less \$5000 between 1990 and 1995, using the sum of damages may generate inconsistency in data over time.

If a candidate treatment county has multiple candidate control counties satisfying all these requirements, we use the synthetic method to calculate the average income of these candidate control counties to construct the control group.¹² If a candidate treatment county does not have any candidate control county satisfying all these requirements, this candidate treatment county is excluded from the treatment group. This can occur for several reasons: (1) all neighboring counties fall into a different metropolitan or non-metropolitan category; (2) all neighboring counties receive a treatment event in the same treatment year; or (3) all neighboring counties receive at least one treatment event in the 15 years before the treatment year, or between the treatment year (yr) and the post-event period (j years after the treatment year: $yr + j$).

The above data generation process is illustrated in Fig. 3. Suppose the treatment event is defined as a relative damage exceeding some cutoff value and the impact duration is 9 years and County A is a non-metropolitan county that receives a treatment event in year 2000. In this case, the treatment year is 2000, the pre-event period is year 1999 and the post-event period is year 2009. Because no other treatment event occurred in this county from 1985 to 1999 (15 years before the treatment year) or from 2001 to 2009 (between the treatment year and the post-event period), this county is a candidate treatment county. We then proceed to construct its counterfactual control group. We start with neighboring counties that share the common border with County A. This allows us to exclude County B1 at the lower-left corner of Fig. 3. We then exclude County B2 and B3 because they are metropolitan counties. We exclude County B4 because it has a treatment event that occurred in the treatment year 2000. We then exclude County B5 because it has a treatment event occurred in 1990 which is within the 15 years prior to the treatment year. We also exclude County B6 because it has a treatment event in 2005, which is between the treatment year (2000) and the post-event period (2009). Consequently, only Counties B7, B8 and B9 are left. For this treatment event in County A in 2000, we construct one counterfactual control county using the average per capita income of these three counties. In this case, we have County A as a county in the treatment group and the counterfactual county in the control group. In cases where we cannot construct a counterfactual county, County A is excluded from the treatment group.

To illustrate how the selection procedure affects the spatial distribution of the treatment group, we plot all selected treatment counties in Fig. 4. It shows the spatial distribution of the treatment counties with cutoff value $k = 250$ and the impact duration $j = 1$. Some candidate treatment counties in Fig. 2 disappear because they do not satisfy some of the requirements.

There are several potential concerns in this selection of the treatment and control group that could lead to underestimate of the significance. Firstly, in cases where we cannot construct a counterfactual county, a candidate treatment county is excluded from the treatment group. Taking hurricanes as an example, counties at the center of hurricanes tend to suffer large damages. However, they are typically excluded from the treatment group because all their neighboring counties are probably stricken by the same hurricane.¹³ The exclusion of these severely affected counties can potentially make the estimates insignificant. Secondly, for a disaster that generates damages in multiple counties, the total damage of all the counties are equally divided among the counties (HVRI 2013). We refer to this procedure as spatial average. This spatial average may overestimate the damage on

¹² As a robustness check, we include all these candidate control counties as the control counties instead of creating the counterfactual using their mean. The results are similar.

¹³ The event of hurricane Katrina is not completely excluded from the treatment because some counties at the periphery of the affected region are included.

<u>County B7:</u> Rural	<u>County B8:</u> Rural	<u>County B9:</u> Rural
<u>County B5:</u> Rural Treatment event in 1990	<u>County A:</u> Rural Treatment event: damage \geq cutoff Treatment year: 2000 Impact duration: 9 years	<u>County B6:</u> Rural Treatment event in 2005
<u>County B2:</u> Urban	<u>County B3:</u> Urban	<u>County B4:</u> Rural Treatment event in 2000
<u>County B1:</u>		

Fig. 3 Construction of the counterfactual county in the control group

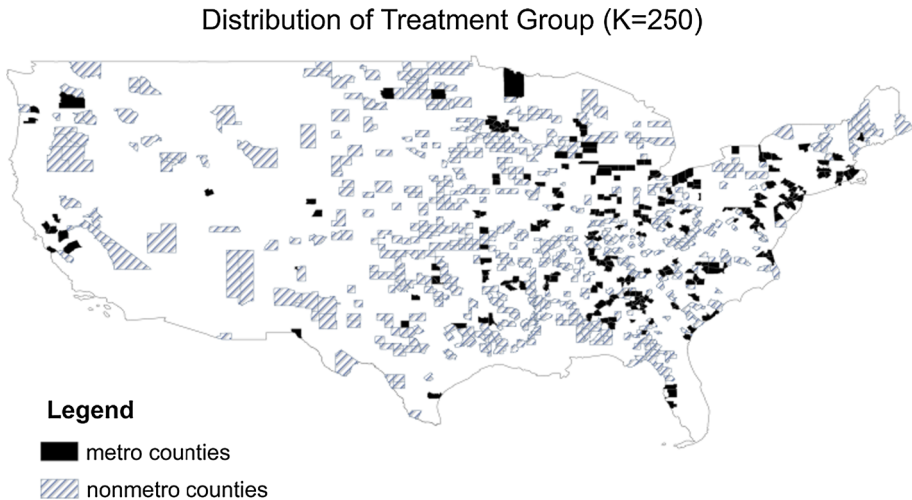


Fig. 4 Distribution of the treatment counties 1990–2012 (cutoff = 250, impact duration = 1)

counties that are marginally affected by the disaster. Comparing a marginally affected county with its neighboring counties not receiving the treatment event tends to generate an insignificant estimate for the impact. In short, in the construction of the treatment group, we tend to exclude counties with severe damages and include counties with marginal damages. In either case, the dataset we generate tends to underestimate the impacts of large natural disasters and leads to a statistically insignificant outcome. Given this tendency to

underestimate the impact, a significant estimated impact is a strong evidence suggesting some real impact of large disasters on income.

Finally, while the use of neighboring counties improves the similarity between the control and the treatment counties, it also introduces the potential spatial spillover effects. Consequently, the estimated difference-in-differences model is not the treatment effect in the common sense. It measures the relative economic performance of disaster-stricken counties as compared to their neighbors and allows the spatial spillover effects to unfold over time. This is an important practical question in itself because the relative economic performance is a key factor determining household/firm location decisions and jurisdictional competition.

3.3 Data characteristics

To investigate how the severity of damage affects income in both the short and long run, we explore various cutoff levels (k) of the relative damage. Here, we present three cutoff values (k) 125, 250 and 500, which roughly correspond to the 92.5th, 95th and 97.5th percentile of the relative damage distribution, respectively.¹⁴

As an example, Table 1 shows the summary statistics of economic variables and the distribution of disaster types for treatment counties across three cutoff values and one impact duration ($j = 1$). Per capita income, average transfer and employment rate in the selected sample are similar to those in the raw data, which include 4249 treated counties and span the years from 1990 to 2012. The top six disaster types by frequency in the selected sample are the same as in the raw data. However, the detailed composition is different. The selected sample contains relatively more incidences of tornado and flooding, and relatively fewer incidences of severe storm/thunder storm and wind.

3.4 Comparison between treatment and control groups before the treatment year

Because we normalize the property damage by average county income in the pre-event period, other things equal, a poor county maybe more likely to be included in the treatment group while a rich county maybe more likely to be included in the control group. To address this concern, we conduct the mean-comparison tests in the pre-event period and report the t-statistics in Table 2. For all cutoff values and impact durations, the mean incomes are not statistically different between the treatment and control groups.

As shown in Fig. 5, the treatment and control groups have almost identical income distribution, and exhibit similar trend of change before the treatment event (see Appendix A in supplementary materials for a more detailed discussion).

4 Results and discussions

4.1 Estimated results

As an example, Table 3 reports the estimation results from Eq. (1) for impact duration $j = 6$. It shows when the cutoff value is either 250 or 500, income of the treatment counties

¹⁴ As a robustness check, we also select sample at other percentiles, but results are consistent.

Table 1 Summary statistics for treated counties from 1990 to 2012 (impact duration = 1)

	<i>K</i> = 125		<i>K</i> = 250		<i>K</i> = 500	
	Raw data	Pre-15	Raw data	Pre-15	Raw data	Pre-15
<i>Socioeconomic characteristics</i>						
Real personal income (\$1000)	14.31	13.33	14.27	13.63	14.44	13.95
Population (1000 persons)	229.71	71.16	252.00	94.96	277.41	127.43
Per capita total transfer (\$)	2436.76	2499.60	2408.58	2565.96	2397.79	2455.84
Per capita personal current transfer Receipts (\$)	2354.79	2417.57	2326.14	2483.70	2316.15	2373.75
Per capita current transfer receipts of individuals from businesses (\$)	34.09	34.49	34.62	34.83	33.49	35.22
Per capita unemployment insurance compensation (\$)	82.80	84.44	83.35	86.42	81.33	87.01
Total full-time and part-time Employment rate (%)	57.96	55.65	58.20	57.38	58.39	58.60
Wage and salary employment rate (%)	47.76	45.41	47.91	47.77	48.01	49.01
Farm employment rate (%)	0.55	1.40	0.51	1.02	0.46	0.89
Non-farm employment rate (%)	57.41	54.25	57.69	56.36	57.93	57.70
Construction sector employment rate (%)	3.24	3.14	3.25	3.13	3.27	3.25
Observations	4249	730	2673	699	1583	559
<i>Disaster type distributions (%)</i>						
Wind	28.05	10.68	27.16	9.25	27.65	10.94
Severe storm/thunder storm	21.29	8.07	21.14	7.46	21.96	6.66
Flooding	16.04	31.3	15.9	34.45	13.43	30.43
Winter weather	8.94	9.19	8.06	7.07	6.46	6.81
Tornado	6.81	21.37	7.08	22.24	7.93	20.92
Hail	5.31	9.32	6.05	9	6.51	9.83
Hurricane/tropical storm	4.93	4.72	6.1	6.17	7.89	7.61
Lightning	2.66	0.5	2.74	0.13	3.07	0
Heat	1.68	0.37	1.51	0.51	0.83	0.32
Drought	1.47	1.49	1.06	0.77	0.55	1.74
Coastal	1.29	0.5	1.51	0.64	1.79	1.11
Wildfire	1	2.11	1.04	1.67	1.38	3.01
Landslide	0.18	0.25	0.22	0.51	0.28	0.16
Fog	0.14	0	0.2	0	0.05	0
Earthquake	0.09	0	0.11	0	0.09	0.32
Avalanche	0.07	0	0.08	0	0.09	0
Tsunami/seiche	0.05	0.12	0.03	0.13	0.05	0.16
Volcano	0	0	0	0	0	0

Raw data refer to the original dataset from SHELDDUS

in the post-event period is significantly lower than that in the control counties. In addition, income is affected by extreme temperature, mostly the fat tail effects from temperature: lower income is associated with colder or hotter temperature.

Table 2 T-statistics of the mean-comparison test of the pre-event income

Impact duration (<i>j</i>)	<i>k</i> = 125	<i>k</i> = 250	<i>k</i> = 500
1	1.1749	0.355	0.498
2	1.1239	0.0366	0.2459
3	1.0526	−0.1785	0.0116
4	1.1559	−0.1646	0.1472
5	0.9716	0.1007	0.3466
6	0.7267	0.1016	0.3641
7	0.6578	0.2468	0.3039
8	0.7323	0.3119	0.2972
9	0.1057	0.3179	0.1075
10	0.2748	0.5015	0.3207

Ho: income in the control and treatment group has the same mean. Ha: income in the control group is less than that in the treatment group. *, ** and *** indicates $p < 0.1$, $p < 0.05$ and $p < 0.01$, respectively

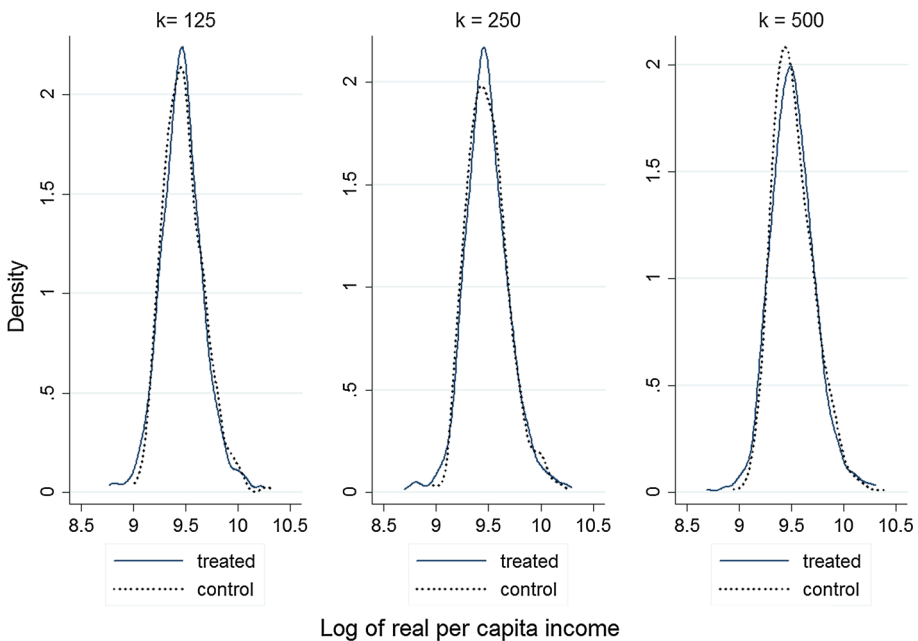


Fig. 5 Distributions of income in the treatment and control groups (impact duration = 1)

We are interested in the relative income difference between the treatment group and the neighboring control group, as well as how such a difference depends on the cutoff values and impact durations. So we report estimates of β_1 in Eq. (1) for each of the three cutoffs ($k = 125, 250, 500$) and each of the ten impact durations ($j = 1, 2, \dots, 10$) in Table 4.

Table 4 presents the main result of this paper: large natural disasters have significantly negative long-run impacts (impact duration $j \geq 5$) on income. Depending on the cutoff value and the impact duration, the difference of the estimated long-run effect between the treatment and control group ranges from -0.008 to -0.02 , corresponding to an average income loss of about \$112 to \$267 in treatment counties as compared to the control counties.

The significant long-run impacts appear to be counterintuitive. However, it reiterates an important point made by Kousky: “different [spatial and temporal] boundaries of analysis can

Table 3 Estimation results of Eq. (1) with impact duration $j = 6$

	$K = 125$	$K = 250$	$K = 500$
Period	0.1269*** (0.0062)	0.1171*** (0.0045)	0.1148*** (0.0041)
DID	-0.0060 (0.0063)	-0.0106** (0.0047)	-0.0080* (0.0046)
Mean temperature	0.0639*** (0.0198)	0.0414*** (0.0129)	0.0257** (0.0109)
Mean temperature, squared	-0.0022* (0.0009)	-0.0010* (0.0006)	-0.0009* (0.0005)
Total precipitation	0.0019 (0.0100)	0.0122 (0.0076)	0.0073 (0.0071)
Total precipitation, squared	-0.0002 (0.0004)	-0.0005* (0.0003)	-0.0004 (0.0003)
Constant	8.9937*** (0.1409)	9.0510*** (0.0996)	9.2871*** (0.0830)
R^2	0.59	0.63	0.68
N	1616	1762	1292

Dependent variable is the log of real per capita income. Standard errors are in parentheses. *, ** and *** indicates $p < 0.1$, $p < 0.05$ and $p < 0.01$, respectively. j in this case equals 6. “DID” is the estimated treatment effects. “Period” is the time effect. N is the number of observations in the panel dataset

generate different results” (Kousky 2014). There could be several mutually non-exclusive explanations for this statistic pattern. Firstly, the short-run increase in the construction demand and capital investment due to property damages helps to smooth the short-run losses in capital stock. Households and firms might have planned for remodeling, reconstruction or capital investment even without the disaster. The occurrence of natural disasters only changes the timing of these investments. Therefore, the short-run increase in rebuilding and reconstruction is achieved at the cost of a decrease in the long-run investment. Secondly, the increase in disaster aids in the short-run diminish over time, as shown in the case of hurricanes (Coffman and Noy 2012; Deryugina 2013), thus the negative impacts become more evident in long run. Thirdly, natural disasters can lead to changes in risk perceptions and induce household/firm relocation among neighboring counties. If the occurrence of large natural disasters makes the treatment counties less desirable as compared to their neighboring counties, the income in the treatment counties will be lower than the control counties after the disaster (Banzhaf and Walsh 2008). Because migration is a slow process, this consequence may only be evident in long run. Finally, the insignificance of the short-run impacts may reflect the limitation of this research: the construction of the treatment and control counties in this research tend to generate insignificant results.

To examine the first two possibilities, we conduct a simple panel analysis using county-level employment and transfer payment¹⁵ data between 1990 and 2012. The panel regression model is written as:

$$\ln q_{iyr} = \alpha_1 \text{Short}_{iyr} + \alpha_2 \text{Middle}_{iyr} + \alpha_3 \text{Long}_{iyr} + \alpha_4 \text{Metro}_{iyr} + \alpha_5 Z_{iyr} + c_i + \gamma_{yr} + \varepsilon_{iyr} \quad (2)$$

¹⁵ Disaster aids are unavailable at the county level.

Table 4 Estimated treatment effects on income for all counties

Impact duration (<i>j</i>)	<i>k</i> = 125			<i>k</i> = 250			<i>k</i> = 500		
	DID	R ²	<i>N</i>	DID	R ²	<i>N</i>	DID	R ²	<i>N</i>
1	-0.0007 (0.0019)	0.35	3172	0.0000 (0.0022)	0.29	3002	-0.0005 (0.0018)	0.38	2348
2	0.0008 (0.0026)	0.37	2760	0.0013 (0.0024)	0.38	2646	0.0007 (0.0023)	0.45	2108
3	0.0017 (0.0031)	0.49	2408	-0.0026 (0.0029)	0.49	2294	-0.0012 (0.0027)	0.54	1832
4	0.0013 (0.0037)	0.57	2224	-0.0031 (0.0034)	0.55	2142	-0.0023 (0.0031)	0.60	1664
5	0.0003 (0.0047)	0.56	1896	-0.0014 (0.0038)	0.61	1942	-0.0022 (0.0036)	0.66	1468
6	-0.0060 (0.0063)	0.59	1616	-0.0106** (0.0047)	0.63	1762	-0.0080* (0.0046)	0.68	1292
7	-0.0073 (0.0076)	0.62	1380	-0.0100* (0.0051)	0.64	1646	-0.0077 (0.0052)	0.69	1204
8	-0.0120 (0.0079)	0.65	1240	-0.0133** (0.0057)	0.67	1446	-0.0115** (0.0056)	0.74	1060
9	-0.0159* (0.0091)	0.66	1088	-0.0147** (0.0061)	0.70	1326	-0.0106* (0.0062)	0.73	984
10	-0.0200** (0.0093)	0.71	940	-0.0162** (0.0068)	0.73	1130	-0.0141** (0.0071)	0.76	836

Dependent variable is the log of real per capita income. *k* refers to the cutoff value. Robust standard errors are in parentheses. *, ** and *** indicates $p < 0.1$, $p < 0.05$ and $p < 0.01$, respectively. *j* is the impact duration from 1 to 10 years. DID is the estimated treatment effects. *N* is the number of observations in the panel dataset

where q_{iyr} is the log employment or log income transfer of county *i* in year yr. Dummy variables $short_{iyr}$, $middle_{iyr}$ and $long_{iyr}$ are used to capture the disaster impacts 1–3, 4–7 and 8–10 years after the treatment event. Dummy variable $metro_{iyr}$ equals one if county *i* is metropolitan and zero otherwise. Z_{iyr} is a vector of weather variables including mean temperature and annual precipitation. c_i and γ_{yr} are the county and year fixed effects, respectively. ε_{iyr} is the error term.

We find both the employment and transfer payment increase in the short-run and decrease in the long-run (see Table 5). These results are consistent to the recent studies on hurricanes (Deryugina 2013), wildfires (Nielsen-Pincus et al. 2014) and tornados (Boustan et al. 2012), suggesting the first two explanations may contribute to the long-run significance in Table 4.

To investigate the last possibility,¹⁶ we apply a panel regression using the raw data:

$$\ln y_{iyr} = \sum_{j=10} \theta_{1j} \text{Treat}_{iyr-j} + \theta_2 \text{Metro}_{iyr} + \theta_3 Z_{iyr} + c_i + \gamma_{yr} + \varepsilon_{iyr} \tag{3}$$

where treat_{iyr} is the dummy variable to show whether a disaster occurs in county *i* in

¹⁶ We do not run regression to investigate the third possibility due to the absence of migration data.

Table 5 Disaster impacts on employment and transfer income in affected counties

	Total employment			Total transfer payment		
	<i>k</i> = 125	<i>k</i> = 250	<i>k</i> = 500	<i>k</i> = 125	<i>k</i> = 250	<i>k</i> = 500
	Short term (1–3 years)	0.0122** (0.0058)	0.0117* (0.0062)	0.0150** (0.0069)	0.0353*** (0.0070)	0.0274*** (0.0081)
Middle term (4–7 years)	0.0033 (0.0066)	0.0073 (0.0076)	-0.0014 (0.0086)	-0.0148** (0.0074)	-0.0027 (0.0090)	-0.0010 (0.0112)
Long term (8–10 years)	-0.0264*** (0.0043)	-0.0317*** (0.0057)	-0.0363*** (0.0070)	-0.0253*** (0.0039)	-0.0293*** (0.0053)	-0.0428*** (0.0063)
Metro	0.0369*** (0.0082)	0.0368*** (0.0082)	0.0369*** (0.0082)	0.0413*** (0.0071)	0.0412*** (0.0071)	0.0412*** (0.0071)
Mean temperature	-0.0073*** (0.0019)	-0.0074*** (0.0019)	-0.0072*** (0.0019)	0.0024 (0.0018)	0.0024 (0.0018)	0.0025 (0.0018)
Mean temperature, squared	0.0003*** (0.0001)	0.0003*** (0.0001)	0.0003*** (0.0001)	-0.0000 (0.0001)	-0.0000 (0.0001)	-0.0000 (0.0001)
Total precipitation	-0.0066*** (0.0012)	-0.0066*** (0.0012)	-0.0066*** (0.0012)	-0.0026* (0.0013)	-0.0026* (0.0013)	-0.0026* (0.0013)
Total precipitation, squared	0.0002*** (0.0001)	0.0002*** (0.0001)	0.0002*** (0.0001)	-0.0000 (0.0001)	-0.0000 (0.0001)	-0.0000 (0.0001)
Constant	9.4508*** (0.0214)	9.4512*** (0.0213)	9.4498*** (0.0213)	10.7253*** (0.0206)	10.7248*** (0.0206)	10.7239*** (0.0206)
<i>R</i> ²	0.3543	0.3542	0.3543	0.8678	0.8677	0.8677
<i>N</i>	70129	70129	70129	69866	69866	69866

Dependent variables are employment and log transfer payment. *k* refers to the cutoff value. Robust standard errors are in parentheses. *, ** and *** indicates *p* < 0.1, *p* < 0.05 and *p* < 0.01, respectively. *R*² is goodness of fit. *N* is the number of observations in the panel dataset

year yr and j is the impact duration from one to ten. e_{iyt} is the error term. The other variables are the same as in Eq. (2).

Equation (3) is run for four different cases as shown in Table 6. In the first case, dummy variable $treat = 1$ if any type of disasters occurs in the county. Otherwise, $Treat = 0$. This is equivalent to the case of cutoff value $k = 0$. The results are reported in the second column. Columns three to five report results of setting cutoff levels $k = 125, 250$ and 500 ,

Table 6 Impacts of disasters on income using the raw data panel sample

	$k = 0$	$k = 125$	$k = 250$	$k = 500$
Treat 1	-0.0009 (0.0010)	-0.0036*** (0.0011)	-0.0011 (0.0013)	0.0013 (0.0017)
Treat 2	-0.0006 (0.0008)	-0.0052*** (0.0012)	-0.0029** (0.0014)	0.0017 (0.0018)
Treat 3	-0.0010 (0.0008)	-0.0049*** (0.0011)	-0.0023* (0.0013)	-0.0003 (0.0018)
Treat 4	-0.0034*** (0.0008)	-0.0033*** (0.0011)	-0.0002 (0.0014)	-0.0002 (0.0018)
Treat 5	-0.0023*** (0.0008)	-0.0042*** (0.0012)	-0.0020 (0.0013)	-0.0030* (0.0017)
Treat 6	-0.0043*** (0.0008)	-0.0043*** (0.0012)	-0.0034** (0.0014)	-0.0045** (0.0018)
Treat 7	-0.0050*** (0.0009)	-0.0071*** (0.0012)	-0.0064*** (0.0014)	-0.0066*** (0.0018)
Treat 8	-0.0076*** (0.0009)	-0.0042*** (0.0012)	-0.0040*** (0.0014)	-0.0059*** (0.0018)
Treat 9	-0.0087*** (0.0010)	-0.0022* (0.0012)	-0.0041*** (0.0014)	-0.0049*** (0.0019)
Treat 10	-0.0059*** (0.0012)	-0.0025** (0.0012)	-0.0038*** (0.0014)	-0.0044** (0.0018)
Metro	0.0156*** (0.0040)	0.0152*** (0.0040)	0.0152*** (0.0040)	0.0152*** (0.0040)
Mean temperature	-0.0008 (0.0015)	-0.0003 (0.0015)	-0.0003 (0.0015)	-0.0004 (0.0015)
Mean temperature, squared	-0.0003*** (0.0001)	-0.0004*** (0.0001)	-0.0004*** (0.0001)	-0.0004*** (0.0001)
Total precipitation	-0.0034*** (0.0010)	-0.0039*** (0.0010)	-0.0039*** (0.0010)	-0.0039*** (0.0010)
Total precipitation, squared	0.0001* (0.0000)	0.0001** (0.0000)	0.0001** (0.0000)	0.0001*** (0.0000)
Constant	9.4685*** (0.0151)	9.4388*** (0.0142)	9.4376*** (0.0142)	9.4379*** (0.0142)
R^2	0.7091	0.7076	0.7073	0.7073
N	70253	70253	70253	70253

Dependent variable is the log per capita income. k refers to the cutoff value. Raw data refers to the panel data without selecting the treatment and control groups. Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. We define the panel data without selecting the treatment and control groups as the raw data

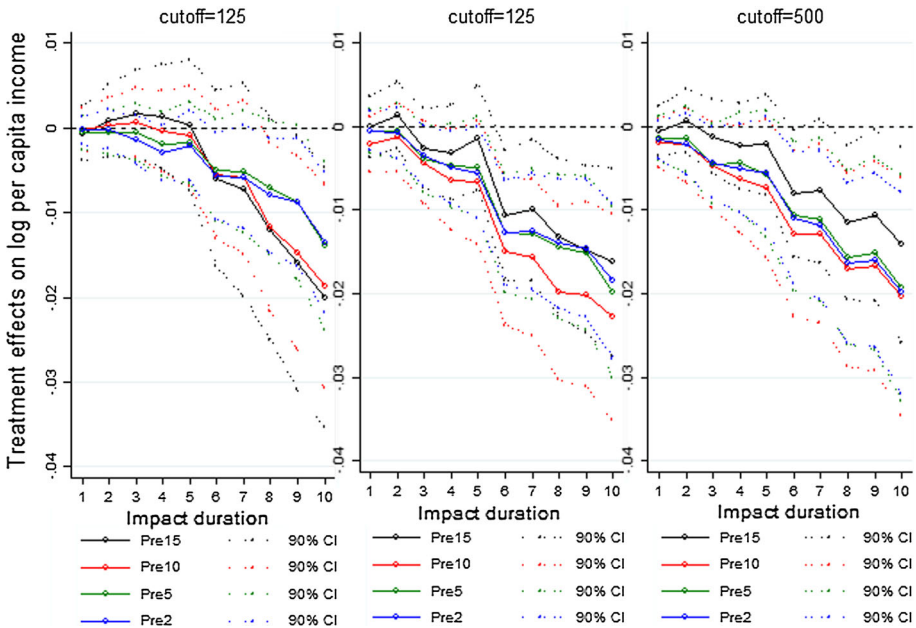


Fig. 6 Relationship between estimated treatment effects and impact durations for three cutoff levels. The x-axis is the impact duration from one to ten, and the y-axis is the estimated treatment effect, together with 90 % confidence interval (CI)

respectively. The only difference with the baseline regression in Table 4 is that the raw data are used in the regression and the assignment of treatment and control groups discussed in Sect. 3.2 is not applied. Results are summarized in Table 6. For all four cases, the long-run impacts of incidences of large disasters are significantly negative. Compared with Table 4, the results are statistically more significant, which is probably due to the serial correlation in the data (see Bertrand et al. 2004). As expected, the insignificant short-run impacts in the baseline model seem to be sensitive to the way the treatment and control groups are constructed.

4.2 Robustness checks

We first test the sensitivity of our results to the disaster-free assumption before the treatment event. In the selection of the treatment and control groups, we impose a constraint that both the treatment and control counties should be free from large disasters in the 15 years before the treatment event. This assumption seems strong but is necessary to control for the potential long-run impacts from past incidences as shown in Table 4 and Table 6. However, it raises concerns about the sensitivity of the results to this particular assumption.

As a robustness check, we relax this assumption and require that both the treatment and control counties should be free of large disasters in the ten, five and two years¹⁷ before the treatment event. Three datasets are constructed based on these three specifications, and

¹⁷ We do not use one year before the treatment event because we expect that large disasters may have immediate impacts on per capita income.

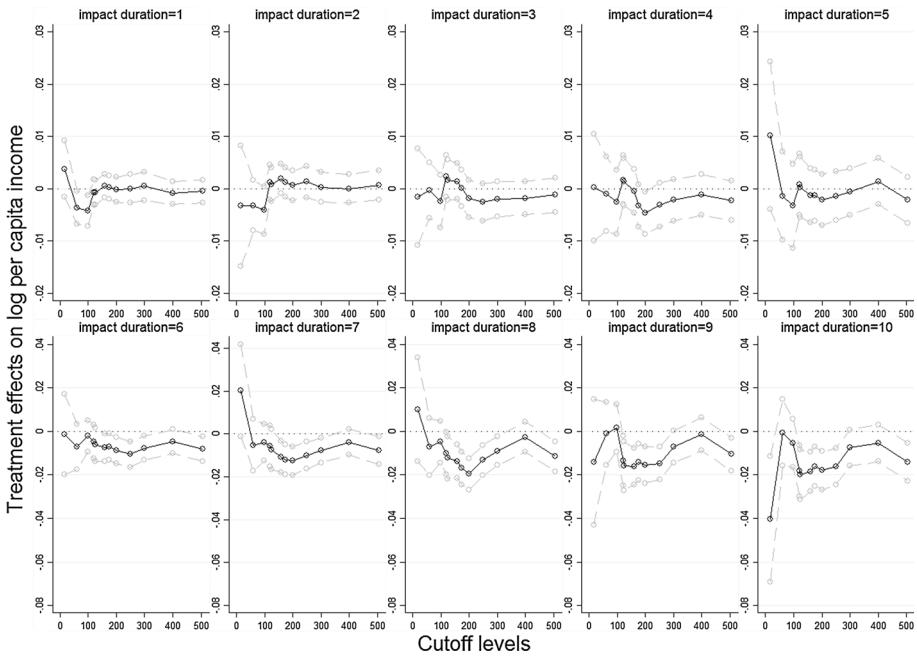


Fig. 7 Relationships of treatment effects and cutoff levels for each impact duration. The cutoff levels are plotted on the x-axis. The estimated treatment effect, together with 90 % confidence interval (CI) are plotted on the y-axis

results are reported in Fig. 6. The estimates from the baseline case with the restriction of 15 years are plotted as the solid line in black. The corresponding dotted lines in black represent the 90 % confidence interval. The estimates and the corresponding confidence intervals for the ten-, five and two-year cases are plotted in red, green and blue, respectively. It is evident that the long-run significance is not sensitive to this specific assumption.

As an extension of the baseline cases with three cutoff values of relative damage, we include another eight cutoff values ranging from 75th to 97.5th quantiles of the relative damage distribution.¹⁸ As shown in Fig. 7, for cutoff values ranging from 125 to 500, significant income differences between the treatment and control groups are observed in long run but not in short run. When the cutoff values are too small or too big, neither the long-run nor the short-run impacts are significant. If the cutoff value is too small, the impact may not be strong enough to generate statistically significant differences. When the cutoff value is too high, the rapid reduction in the number of observations results in the statistical insignificance.

As an alternative specification, we use the detrended income as the dependent variable in Eq. (1). As shown in Table 7, results are consistent with the baseline results in Table 4. We also split out sample by metropolitan and non-metropolitan counties and find similar results (see Appendix B in supplementary materials).

¹⁸ Further decrease in the cutoff value results in insufficient number of observations.

Table 7 Estimated treatment effects on detrended income for all counties

Impact duration (<i>j</i>)	<i>k</i> = 125			<i>k</i> = 250			<i>k</i> = 500		
	DID	<i>R</i> ²	<i>N</i>	DID	<i>R</i> ²	<i>N</i>	DID	<i>R</i> ²	<i>N</i>
1	10.4598 (25.3467)	0.05	3176	−10.0323 (24.9813)	0.03	3004	−4.4423 (24.4764)	0.05	2348
2	40.9818 (31.4371)	0.05	2764	26.9333 (32.8853)	0.03	2648	29.4639 (30.5806)	0.05	2108
3	44.0612 (42.1339)	0.07	2412	−37.0958 (39.3939)	0.05	2296	−1.5188 (38.5933)	0.09	1832
4	19.2608 (48.6368)	0.10	2228	−46.7145 (47.6136)	0.07	2144	−28.2804 (46.1216)	0.05	1664
5	26.5962 (62.2570)	0.09	1900	−20.0934 (56.0071)	0.09	1944	−34.1774 (54.1544)	0.06	1468
6	−21.7961 (83.6059)	0.13	1620	−162.5050** (71.5423)	0.10	1764	−121.4715* (66.1435)	0.08	1292
7	13.4629 (97.5030)	0.13	1384	−144.2977* (74.0433)	0.08	1648	−123.7736 (76.8617)	0.10	1204
8	−71.0171 (95.8273)	0.16	1244	−184.2279** (84.0028)	0.09	1448	−165.4711** (81.5183)	0.12	1060
9	−99.3366 (127.0240)	0.15	1092	−213.1511** (95.0187)	0.14	1328	−155.9836 (98.3059)	0.11	984
10	−180.6028 (124.3712)	0.20	944	−227.6219** (107.7250)	0.15	1132	−164.3584 (102.5723)	0.12	836

Robust standard errors are in parentheses. *, ** and *** indicates $p < 0.1$, $p < 0.05$ and $p < 0.01$, respectively. *j* is the impact duration from 1 to 10 years. DID is the estimated treatment effects. *N* is the number of observations in the panel dataset

5 Conclusion

In this paper, we examine the impact of large natural disasters occurred between 1990 and 2012 on county per capita income using the SHEL DUS dataset that covers 18 types of natural hazards. In contrast to existing regional studies that typically focus on one specific type of disasters, we investigate the regional economic impacts of multiple types of natural disasters.

Using a difference-in-differences model with fixed effects and controlling for serial correlation in the county per capita income data, we find that comparing with neighboring counties, large natural disasters have significant long-run impact, and the difference of the real per capita income falls in a range between \$112 and \$267. Recognizing these differences and developing location-based emergency plans are important for effective disaster assistance programs. Results from this paper also suggest that we probably need to reform the existing disaster relief programs in order to mitigate the long-run impacts.

References

Anttila-Hughes JK, Hsiang SM (2013) Destruction, disinvestment, and death: economic and human losses following environmental disaster. Available at SSRN 2220501

- Banzhaf HS, Walsh RP (2008) Do people vote with their feet? An empirical test of Tiebout's mechanism. *Am Econ Rev* 98:843–863
- Belasen AR, Dai C (2014) When oceans attack: assessing the impact of hurricanes on localized taxable sales. *Ann Reg Sci* 52:325–342
- Belasen AR, Polachek SW (2008) How hurricanes affect wages and employment in local labor markets. *Am Econ Rev* 98:49–53
- Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Q J Econ* 119:249–275
- Boustan LP, Kahn ME, Rhode PW (2012) Moving to higher ground: migration response to natural disasters in the early twentieth century. *Am Econ Rev* 102:238–244
- Brown C, Meeks R, Ghile Y, Hunu K (2013) Is water security necessary? An empirical analysis of the effects of climate hazards on national-level economic growth. *Philos Trans Ser A Math Phys Eng Sci* 371:20120416
- Cavallo E, Noy I (2011) Natural disasters and the economy—a survey. *Int Rev Environ Res Econ* 5:63–102
- Cavallo E, Galiani S, Noy I, Pantano J (2013) Catastrophic natural disasters and economic growth. *Rev Econ Stat* 95:1549–1561
- Coffman M, Noy I (2012) Hurricane Iniki: measuring the long-term economic impact of a natural disaster using synthetic control. *Environ Dev Econ* 17:187–205
- Dell M, Jones BF, Olken BA (2012) Temperature shocks and economic growth: Evidence from the last half century. *Am Econ J Macroeconomics*:66–95
- Deryugina T (2013) The role of transfer payments in mitigating shocks: evidence from the impact of hurricanes. Available at SSRN
- Horowitz JK (2009) The income–temperature relationship in a cross-section of countries and its implications for predicting the effects of global warming. *Environ Res Econ* 44:475–493
- Husby TG, Groot HL, Hofkes MW, Dröes MI (2014) Do floods have permanent effects? Evidence from the Netherlands. *J Reg Sci* 54:355–377
- HVRI (2013) The spatial hazard events and losses database for the United States, version 12.0. Hazards and Vulnerability Research Institute. Columbia, SC: University of South Carolina
- Kousky C (2014) Informing climate adaptation: a review of the economic costs of natural disasters. *Energy Econ* 46:576–592
- Misiewicz J (2011) Fat-tailed distributions: data, diagnostics, and dependence. *RFF*
- Nielsen-Pincus M, Moseley C, Gebert K (2014) Job growth and loss across sectors and time in the western US: the impact of large wildfires. *For Policy Econ* 38:199–206
- Strobl E (2011) The economic growth impact of hurricanes: evidence from US coastal counties. *Rev Econ Stat* 93:575–589
- Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46:234–240
- Vu TB, Hammes D (2010) Dustbowls and high water, the economic impact of natural disasters in China. *Asia Pac J Soc Sci* 1:122–132
- Xiao Y (2011) Local economic impacts of natural disasters. *J Reg Sci* 51:804–820