17

ISBN: 978-602-14328-0-8

ICAS-2013

PROCEEDING

The International Conference On Applied Statistics

September, 16-18th 2013
Universitas Padjadjaran
Kampus Jatinangor

ISBN: 978-602-14328-0-8

ICAS-2013

PROCEEDING

The International Conference On Applied Statistics

September, 16-18th 2013
Universitas Padjadjaran
Kampus Jatinangor

ICAS-2013

The International Conference on Applied Statistics

PROCEEDING

ISBN: 978-602-14328-0-8

Editors:
Septiadi Padmadisastra, Ph.D
Budhi Handoko, S.Si, M.Si
Anindya Apriliyanti P, S.Si, M.Si
Defi Yusti F, S.Si, M.Si

September, 16-18th 2013
Department of Statistics
Faculty of Mathematics and Natural Sciences
Universitas Padjadjaran

Preface

Welcome to the Proceedings of the International Conference on Applied Statistics (ICAS 2013). This event was held at September, 16-18th 2013 in the beautiful Campus of Universitas Padjadjaran, Jatinangor-Sumedang, West Java Province, Indonesia. This Conference is organized by **Department of Statistics, Faculty of Mathematics and Natural Science, Universitas Padjadjaran** in cooperation with **Statistical Forum for Higher Education** (FORSTAT) Indonesia.

The ICAS 2013 Conference was attended by 12 invited speakers, 60 researchers and academicians from 8 countries who were share their expertise and find out about the latest research on applied statistics. A special thanks goes to all speakers and presenters as well as to the session chairpersons, who drove all the conference sessions on the right tract, keeping them in time while permitting enriching discussions.

Finally, we hope that these proceedings will serve as a valuable reference for security researcher and academicians.

Best Wishes,

Dr. Toni Toharudin, M.Sc.
Chair of Organizing Committee

Table of Contents

MULTIVARIATE <M> & TIME SERIES <TS>

SPATIAL <S> & BIOSTATISTICS <B>

SOCIOMETRICS <Sc> & STATISTICAL EDUCATION <SEd>

STATISTICAL THEORY <ST>

THE SYNTHETIC REGRESSION METHOD: HOW THE INDIAN OCEAN TSUNAMI AFFECTS GROWTH TRAJECTORIES

SianaHalim[1], Inggrid[2], Ridhotama Shanti D. Ottemoesoe[2]

[1]Industrial Engineering Department, Petra Christian University
Surabaya, Indonesia
halim@petra.ac.id

[2]Business and Management Department, Petra Christian University
Surabaya, Indonesia
inggrid@petra.ac.id, shanti@petra.ac.id

Abstract

In recent works, scholars have widely used the synthetic control method (SCM) for analyzing the effect of a catastrophic disaster on economic growth. Unlike the traditional SCM, we propose the so-called synthetic regression method to choose a suitable synthetic control group by employing a comparing series procedure. We then use the selected group to estimate our outcomes of interest. We apply the synthetic regression method to examine the causal effects of the 2004 Asian tsunami on economic growth in Aceh, the Indonesian province severely hit by the tsunami. Our results show that the tsunami has an unintended effect on Aceh's economic growth performance. The findings of the paper lead us to a better method for analyzing the growth trajectories, particularly for Aceh after the 2004 tsunami

Keywords: Natural disasters, Economic growth, Synthetic control method, Comparing series, Synthetic regression method.

1. INTRODUCTION

An influential stand of literature on the economics of natural disasters has been devoted to assess the economic effects of such events. The standard taxonomy classifies damages caused by disasters into three classes: direct damages that entail the loss of productive investments (e.g. destruction of fixed assets and inventories), indirect damages that encompass the decline in the production of goods and services, and secondary effects that involve the evolution of macroeconomic variables (Peeling *et al*., 2002; ECLAC, 2003). This paper assesses the change in the growth trajectories of Aceh, the Indonesian province exposed to the 2004 tsunami.

In practice, we adopt an experimental design in which the catastrophic natural disaster or the treatment is supposed to generate a jump or an interrupted in the time series data. Previous research studied the conditions before and after that interruption in different methods. Halim and Jiang (2013) used the interrupted time series analysis with a segmented regression method for analyzing the effect of operation 24 hours on reducing collision in the City of Edmonton. Hamed *et al.* (1999) employed a diffusion model with jumps to investigate the impact of the Gulf crisis on traffic collisions in Jordan. The interrupted time series and diffusion models will produce satisfactory results if we are able to compare the before and after conditions for a single time series.

On the contrary, the synthetic control method (SCM) offers us the opportunity to select an appropriate control group based on multiple time series, and it allows us to construct the synthetic control unit from a group of untreated units. However, the SCM has several drawbacks. For instance, in some cases, it may not even be possible to obtain a weighted combination of untreated units such that some restrictions for constructing the SCM are satisfied. This condition will lead to a poor fit. Even if there is a synthetic control unit that provides a good fit for the treated unit, interpolation biases may be large if the simple linear model does not hold over the entire set of regions in any particular sample (Abadie *et al.*, 2010).

Therefore, in this paper, we propose the so-called synthetic regression method. Basically, it is carried out in two steps. First, we use the comparing series method as Franke and Halim (2007) to select a suitable control group. In our case, we need to choose a proper synthetic control group for the Aceh Province. In the second stage, we use the selected group to estimate economic growth that would have been observed for Aceh in the absence of the disaster. The economic growth is estimated using linear regression.

2. METHODOLOGY

In this section, we give an overview of the two methods for studying the evolution of Aceh's economic growth after the tsunami. They are the synthetic control method and the synthetic regression method.

2.1. Synthetic Control Method

We assume that the Indian Ocean tsunami in 2004 is an exogenous event, and it produces a sizeable effect on the exposed province. The SCM constructs a counterfactual group whose outcomes are compared to the Aceh Province. The synthetic group is a weighted combination of the unexposed provinces that resemble Aceh in economic characteristics before the disaster took place. After the occurrence of the disaster, we estimate the counterfactual situation of Aceh in the absence of the tsunami through the outcomes of the synthetic group. We formalize this concept as follows.

We observe 26 provinces in Indonesia for the period $t = 1994, \dots, 2004, \dots, 2011$. Let $i = 1$ be the Aceh Province, and $i = 2, \dots, 26$ be the other provinces. Here, we let $T_0 = 2004$ be the year when tsunami struck Aceh and Y_{it} be the outcome variable which evaluates the impact of the tsunami on province i at the time t. Additionally, Y_{it}^I is the outcome variable that refers to the tsunami and Y_{it}^N is the outcome variable when the tsunami did not occur. Our goal is to estimate the effect of the tsunami on the economics of the Aceh Province during 2005-2011. This effect is defined as the difference between the outcome variable that refers to the tsunami and when it did not occur, or:

$$\alpha_{1t} = Y_{1t}^I - Y_{1t}^N, t = 2005, \dots, 2011 \quad (1)$$

However, the Y_{1t}^N in those periods is unobserved. Therefore, this outcome will be estimated synthetically using the SCM.

Following Abadie *et al.* (2010), to construct the synthetic control, we define a (25x1) vector of weights $W = (w_2, \dots, w_{26})'$ such that $w_i \geq 0$ for $i = 2, \dots, 26$ and $\sum_{i=2}^{26} w_i = 1$. Abadie and Gardeazabal (2003) and Abadie *et al.* (2010) choose $W^* = \sum_{i=2}^{26} w_i^*$ such that

$$\alpha_{1t} = Y_{1t} - \sum_{i=2}^{26} w_i^* Y_{it}, t = 2005, \dots, 2011 \quad (2)$$

For calculating the synthetic control methods, we use R.3.0 with Synth-R. packages (Abadie *et al.*, 2011).

2.2. Synthetic Regression

The synthetic regression method is a simple linear regression model: $= g(X) + \varepsilon$, where Y is the dependent variable, in this case is Aceh's economic growth, $g(.)$ is a linear function and X is the independent variables. Those independent variables are the synthetic regions whose GRDP/Capital is similar to the growth rate of Aceh.

The synthetic regions are chosen by using a comparing series procedure, as explained in the following section.

Comparing Series

We consider two series of the form (Franke and Halim, 2007)

$$Y_i = m_I(x_i) + \varepsilon_i; Y_i = m_{II}(x_i) + \varepsilon_i; x_i = \frac{i}{n}; i = 0, \dots, n \quad (3)$$

where $\varepsilon_i, \varepsilon_i, i = 0, \dots, n$ are independent with zero mean and finite variance, $var\,(\varepsilon_i)$=$var(\varepsilon_i) = \sigma^2(x_i)$. m_I, m_{II} are the general functions in which m_I represent the series for Aceh and m_{II} represent the series other than Aceh. We want to find a synthetic region which has similar economic characteristics as Aceh, i.e., we want to test the hypothesis $H_o\colon m_I = m_{II}$. m_I, m_{II} are arbitrary functions. Because x_i is equidistant, we can estimate m_I, m_{II} by using the Priestley-Chao (1972) kernel estimator as:

$$m_I(x,h) = \frac{1}{n+1}\sum_{i=0}^{n} K_h(x-x_i)Y_i;\; m_{II}(x,h) = \frac{1}{n+1}\sum_{i=0}^{n} K_h(x-x_i)Y_i \qquad (4)$$

where $K_h(.) = h^{-1}K(./h)$ denotes a rescaled kernel function. K is a probability density symmetric around zero, $h > 0$ is the bandwidth to control the smoothness of the estimated functions.

To perform the test, we measure the distance between the estimated functions $m_I(x,h)$ and $m_{II}(x,h)$. We reject the null hypothesis if that distance is too large. The distance itself is defined as:

$$T_n = n\sqrt{h}\int\left(m_{II}(x,h) - m_I(x,h)\right)^2 dx \approx \sqrt{h}\sum_{i=0}^{n}\left(m_{II}(x,h) - m_I(x,h)\right)^2 \qquad (5)$$

In the same vein, we reject the null hypothesis at a level of α if $T_n > C_\alpha$, where C_α is approximated by using a bootstrap procedure.

The bootstrap procedure is started by estimating the residuals:

$$\varepsilon_i = Y_i - m_I(x,h);\; \varepsilon_i = Y_i - m_{II}(x,h);\; i = 0, \dots, n \qquad (6)$$

We then generate the bootstrap residuals $\varepsilon_i^*, \varepsilon_i^*, i = 0, \dots, n$ and consider two bootstrap series which are constructed as:

$$Y_i^* = m_I(x,h) + \varepsilon_i^*;\; Y_i^* = m_{II}(x,h) + \varepsilon_i^*;\; i = 0, \dots, n \qquad (7)$$

The bootstrap for the kernel estimators is then calculated by using:

$$m_I^*(x,h) = \frac{1}{n+1}\sum_{i=0}^{n} K_h(x-x_i)Y_i^*;\; m_{II}^*(x,h) = \frac{1}{n+1}\sum_{i=0}^{n} K_h(x-x_i)Y_i^* \qquad (8)$$

The bootstrap test statistic can be constructed as:

$$T_n^* = nh^{1/2}\int\left(m_{II}^*(x,h) - m_I^*(x,h)\right)^2 dx \qquad (9)$$

To get a Monte Carlo approximation $C_{\alpha B}^*$ of quantile C_α^*, we generate the realization of T_n^* repeatedly as many as B (e.g. B = 10,000).

3. RESULTS AND DISCUSSION

We study the causal effect of catastrophic natural disasters on the short- and medium-term economic growth. We take a direct case of the 2004 Asian tsunami in Indonesia. One day after Christmas in 2004 at 00.59 GMT, a 9.0 Richter-magnitude earthquake struck the west coast of Sumatra, a major island in the western part of Indonesia. It subsequently generated ferocious tsunami waves in the Aceh Province of Sumatra. This catastrophic disaster is considered as one of the worst tsunamis in human history. According to the official statistics, the death toll of the Aceh tsunami was more than 165,000 people and over a half million people were displaced. The total estimate of economic damages and losses was approximately US$ 4.5 billion corresponding to 97.4% of Aceh's gross regional domestic product (GRDP) in 2003 (Athukorala and Resosudarmo, 2005; Athukorala, 2012).

3.1. Data

The data set was taken from the Indonesian Central Bureau of Statistics (BPS) and includes 26 Indonesian provinces[1]. The data set is annual and covers the period 1994-2011. All nominal variables are converted in 2000 prices. The description of the variables is given below.

The outcome variable is economic growth. It is measured by the change in gross regional domestic product (GRDP) per capita. The predictors of economic growth are:

- The sectoral value added consists of nine major sectors in the economy, that is, agriculture, mining and quarrying, manufacturing, construction, utilities, trade, hotel, and restaurant, transportation and communication, finance, and services. They are also in growth rate.
- The variable of investment is defined as the growth rate of gross fixed capital formation.
- The growth rate of government consumption.
- The proxies for human capital are of the enrollment numbers of primary school, junior high, senior high school, and university. They are expressed in logarithms. The other variables are adult literacy rate and years of schooling.
- Population density is measured as total population divided by land area in kilometer square.

[1] Prior to the fall of the New Order Regime in 1998, Indonesia had 27 provinces. It turned to be 26 provinces after East Timor gained its independence in 1999. The remaining provinces have proliferated today, becoming 34 provinces. However, to maintain consistency, our analysis still uses 26 provinces.

3.2. Results

3.2.1 Using Synthetic Regression Method

We first compare the growth rate of the gross regional domestic product per capita (or GRDP/Cap) of each of the 25 provinces to the Aceh Province by using the comparing series method. We find that Riau, South Sumatra, Jogjakarta, East Nusa Tenggara, Central Kalimantan, and East Kalimantan have similar GRDP/Cap characteristics as Aceh.

We then formulate a linear regression model with the GRDP/Cap of Aceh as the dependent variable and the GRDP/Cap of the selected regions as the independent variables. The estimated coefficients show that the GRDP/Cap of Aceh closely resembles Riau, East Nusa Tenggara, and East Kalimantan.

Table 1. The estimated regression coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-0.1134	0.02266	-5.007	0.000398 ***
Riau	-1.3287	0.64393	-2.063	0.063496
South Sumatra	-0.3029	0.29518	-1.026	0.326845
Jogjakarta	-0.8098	0.50934	-1.59	0.140176
East Nusa Tenggara	3.6155	0.90144	4.011	0.002048 **
Central Kalimantan	0.53476	0.54444	0.982	0.347097
East Kalimantan	-0.9485	0.44009	-2.155	0.054147

Using those three regions, we formulate the second linear model and plot the prediction of that model as the GRDP/Cap of the synthetic region against the GRDP/Cap of Aceh (Fig. 1). From that figure, we can see that the GRDP/Cap of Aceh and the synthetic group is very similar. The figure also confirms that after 2004 (the year of the Asian tsunami), the GRDP/Cap of the synthetic regions is higher than the Aceh Province. This finding lends support to the hypothesis that the catastrophic tsunami negatively affects the GRDP/Cap growth.

Table 2. The estimated regression coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-0.09590	0.02054	-4.668	0.000362 ***
Riau	-1.15996	0.48452	-2.394	0.031223*
East Nusa Tenggara	2.42959	0.54078	4.493	0.000506 ***
East Kalimantan	-0.91195	0.40563	-2.248	0.041192 *

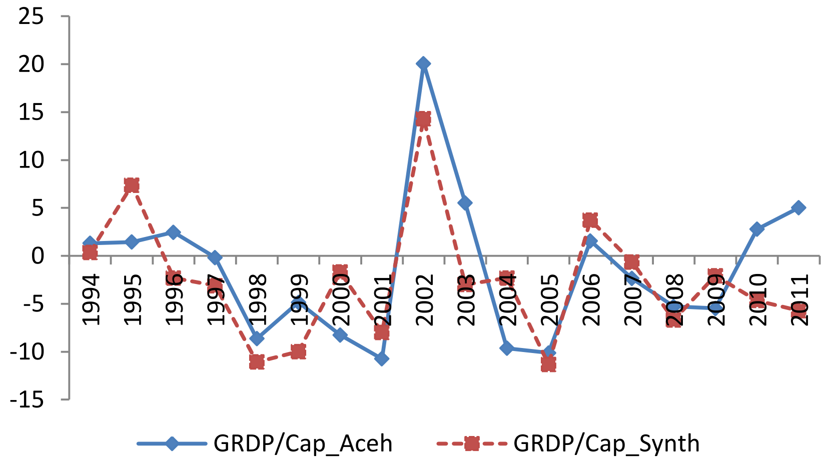


Fig. 1 GRDP/Cap: Aceh vs Synthetic Regions

3.2.2 Using Synthetic Control Method

According to the synthetic control method (SCM), the Aceh's growth trajectory prior to the tsunami is best represented by a combination of DKI Jakarta, Riau, East Nusa Tenggara, Central Kalimantan, and East Kalimantan, the provinces with positive weights (Table 3). It is important to note that these synthetic control provinces are also captured in our proposed method, with the exception of DKI Jakarta. However, we cast some doubts on the SCM results since the GRDP/Cap for the synthetic regions is above the GDP/Cap for Aceh in most of the years, except in the year 2002 (Fig. 2, left). In addition, the GRDP/Cap gap of Aceh and the synthetic provinces is also negative, apart from the year 2002 (Fig. 2, right).

Table 3. Synthetic Weight for Aceh

Province	Weight	Province	Weight
North Sumatra	0	West Nusa Tenggara	0
West Sumatra	0	East Nusa Tenggara	0.239
Riau	0.276	West Kalimantan	0
Jambi	0	Central Kalimantan	0.142
South Sumatra	0	South Kalimantan	0
Bengkulu	0	East Kalimantan	0.235
Lampung	0	North Sulawesi	0.001
DKI Jakarta	0.106	Central Sulawesi	0
West Java	0	South Sulawesi	0
Central Java	0	Maluku	0
Jogjakarta	0	Papua	0
East Java	0		
Bali	0		

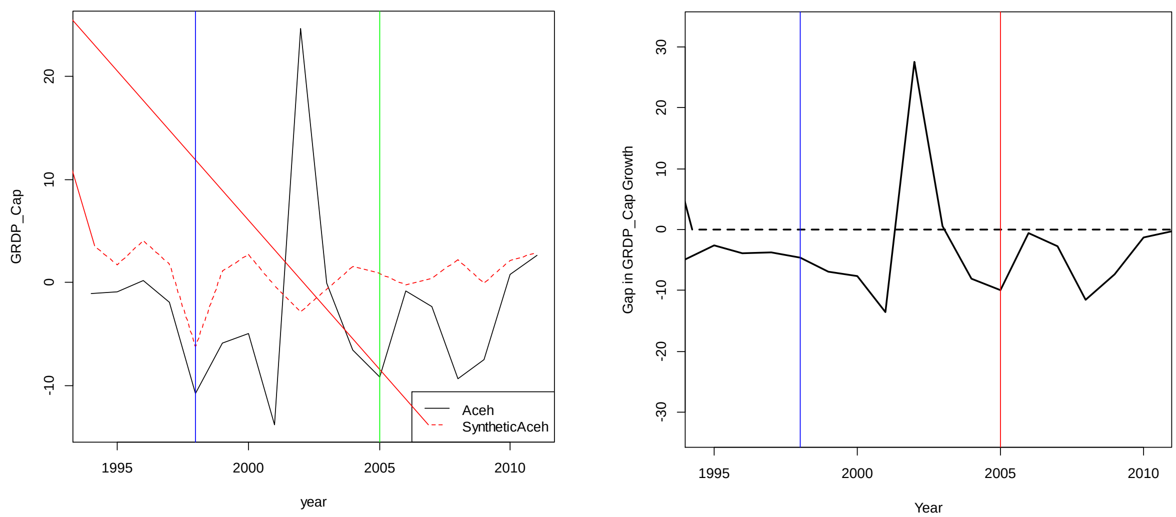
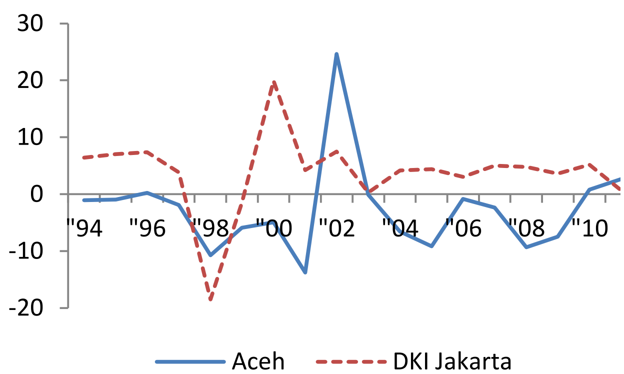


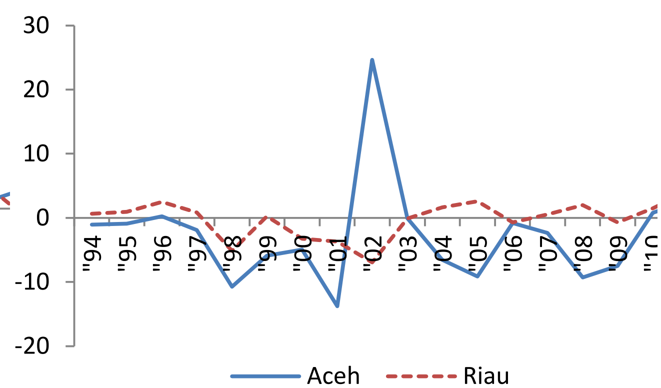
Fig. 2 GRDP/Cap: Aceh vs Synthetic Regions (left), GRDP/Cap Gap: Aceh vs Synthetic Regions (right)

3.3 Discussion

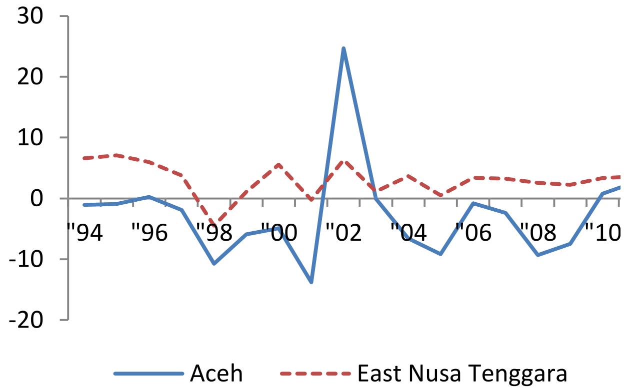
It is well known that, in some instances, the synthetic control method (SCM) may result in poor model fit. Under this condition, it is recommended for not using the SCM (Abadie *et al.*, 2010). In our case, the SCM does not give a reasonable choice of the synthetic regions, particularly for choosing DKI Jakarta (the capital of Indonesia) as one of the synthetic regions for Aceh in terms of the GRDP/Cap growth. The graphical plot of the GRDP/Cap yearly basis between Aceh ***vis-à-vis*** DKI Jakarta shows that most of the time, the GRDP/Cap of DKI Jakarta is above Aceh, except in the year 1998 and 2002 (Fig. 3). With respect to the unusual trend in 1998, we argue that Aceh's economy grew stronger than the economy of DKI Jakarta even though the two provinces experienced worse economic downturns as a result of the Asian financial crisis.[2] This argument is fairly acceptable since the crisis led to the deepest contractions in the industrial and service sectors, the two sectors that dominated DKI Jakarta's economy, whereas the main pillars of Aceh's economy (i.e. the agricultural and oil and gas sectors) were less affected. As for the year 2002, high price levels for oil and gas and a substantial amount of Special Autonomy Fund (Dana Otsus) from the central government partly contributed to Aceh's remarkable economic growth.

On the other side, both Aceh and the three synthetic regions that are chosen by our proposed method seem to follow a similar economic growth path from year to year, except for the year of 2002 (Fig. 3).[3] Therefore, we highlight that our chosen method gives more reasonable results as compared to the SCM. From the theoretical part, our findings suggest that the tsunami disaster has a negative impact on regional economic growth. This evidence is consistent with previous studies that employ the standard SCM, such as Noy (2009), Coffman and Noy (2012), and Cavallo *et al*. (forthcoming).





[2] In 1998, the Asian financial crisis severely hit Indonesia. It triggered a big riot in Jakarta and forced the Indonesian President at that time (the late President Soeharto) to end his 32 years presidency in Indonesia.
[3] With regard to the rapid growth of Aceh in 2002, we just provide two plausible explanations in the previous paragraph.



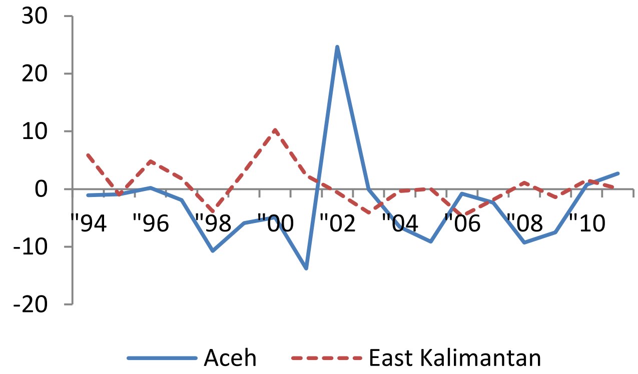


Fig 3.GRDP/Cap: Aceh vs Synthetic Regions

4. CONCLUSION

We show that the synthetic regression method provides more convincing results as compared to the traditional synthetic control method. As the heart of the synthetic regression method, the comparing series procedures are able to choose the synthetic control provinces that best resemble Aceh's economic growth trajectory. Our results are in the right direction, but we have not finished yet. Finding a well-defined approach to determine the weight of the synthetic regions instead of using simple linear regression models opens up new avenues for our future work.

ACKNOWLEDGMENT

The authors gratefully acknowledge financial support for this research from the Indonesian Directorate General of Higher Education (DIKTI).

5. REFERENCES

Abadie, A., Diamond, A. &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*,Volume 105**,** 493–505.

Abadie, A., Diamond, A. &Hainmueller, J. (2011). Synth: An R Package for Synthetic Control Methods in Comparative Case Studies. *Journal of Statistical Software*,Volume42**,** 1-17.

Athukorala, P.-C.(2012). Indian Ocean Tsunami: Disaster, Generosity, and Recovery. *Asian Economic Journal*,Volume 26**,** 211-231.

Athukorala, P.-C.&Resosudarmo, B. P. (2005). The Indian Ocean Tsunami: Economic Impact, Disaster Management, and Lessons. *Asian Economic Papers*,Volume 4**,** 1-39.

Cavallo, E. A., Galiani, S., Noy, I. & Pantano, J. forthcoming. Catastrophic Natural Disasters and Economic Growth. *The Review of Economics and Statistics*.

Coeffman, M. & Noy, I. (2012). Hurricane Iniki: Measuring the Long-term Economic Impact of a Natural Disaster Using Synthetic Control. *Environment and Development Economics,* Volume 17, 187-205.

Eclac(2003).*Handbook for Estimating the Socio-Economic and Environmental Effects of Disasters*, United Nations Economic Commission for Latin America and the Caribbean.

Franke, J., &Halim, S. (2007). Wild Bootstrap Tests: Regression Models for Comparing Signals and Images, *IEEE Signal Processing Magazine* July 2007, 31-37.

Halim, S. & Jiang, H. (2013). The Effect of Operation 24 Hours on Reducing Collision in the City of Edmonton, *Accident Analysis and Prevention*, Volume 58, 106-114.

Hamed, M.M, Al-Fideh, B.M., Al-Sharif, M.M.(1999). Traffic Accidents under the Effect of the Gulf Crisis, *Safety Science*, Volume 33, 59-68.

Noy, I. 2009.The Macroeconomic Consequences of Disasters.*Journal of Development Economics,* Volume 88**,** 221-231.

Pelling, M., Ozerdem, A. &Barakat, S. (2002). The Macro-Economic Impact of Disasters.*Progress in Development Studies*,Volume 2**,** 283-305.

Priestley, M.B. & Chao, M.T. (1972). Nonparametric Function Fitting, *Journal of Royal Statistics Society, Series* B, Volume 34, 385-392.