Does schooling reduce juvenile delinquency? Evidence from a natural experiment in Japan*

Yu AOKI†

March 2010

[Preliminary Version]

Abstract

Individual returns of schooling have been examined intensively but social returns of schooling have received less attention. If schooling yields not only individual returns but also social returns – such as a reduction in crime – then the rationale for policies which encourage individual investments in schooling is strengthened. In this paper, we explore the effect of schooling on juvenile delinquency. To identify the effect of schooling, we use exogenous variation in schooling caused by a policy intervention in specific cities and towns after the Great Hanshin-Awaji Earthquake, which occurred in the Midwestern part of Japan in 1995. We treat this as a natural experiment. Using the Within Group Instrumental Variable estimator, we address an endogeneity problem of schooling caused by simultaneity and unobserved heterogeneity. Our results indicate that higher school attainment significantly reduces crime committed by youth regardless of crime categories, i.e., reduces both violent crime and property crime. This negative impact of schooling on crimes supports theoretical predictions that higher educational attainment reduces the participation in criminal activities through a higher opportunity cost of committing crimes.

Keywords: schooling, crime, natural experiment

JELClassification: H52, I28, K42

---

*Acknowledgment: I am grateful to Professor Wiji Arulampalam and Professor Robin Naylor to their expert supervisions and continuous supports. All remaining errors are my responsibility.

†Department of Economics, University of Warwick, Coventry, CV4 7AL, United Kingdom, Tel: +44-(0)24-765-28413, E-mail: Y.Aoki@warwick.ac.uk

1 Introduction

To date, individual returns of schooling have been examined intensively but social returns of schooling have received less attention. If schooling yields not only individual returns but also social returns – such as a reduction in crime – then the rationale for policies which encourage individual investments in schooling is strengthened. Furthermore, crime prevention by, for example, increasing an individual's educational level may cost less than combating crime after it has occurred.

This paper attempts to draw inferences on the causal relationship between schooling and juvenile delinquency using an instrumental variable naturally occurring in a natural experiment in Japan. If a causal relationship rather than just mere correlation is revealed, the Government can target a policy to reduce juvenile delinquency. Using an instrumental variable to obtain a consistent estimator is important, since the causal variable in focus, a schooling indicator, is potentially endogenous. Specifically, individuals' unobserved heterogeneity such as a variation in learning ability or criminal ability may affect individuals' criminal participation decisions. Moreover, the more an individual spends time committing crime, the less the individual may want to spend time to attain education due to the increased probability of being arrested before the benefits of education can be reaped (reverse causality).

In this paper, we focus on the governmental support granted to students, teachers and schools in specific cities and towns after the Great Hanshin-Awaji earthquake occurred in the Midwestern part of Japan in 1995. The support was granted to students, teachers and schools in the disaster area regardless of their abilities or performances. Hence, the support would have desirable properties as an instrument: they are not expected to be correlated with unobserved heterogeneity. Further, the support would not have a direct effect on crime itself. Given that most existing literature does not address the potential endogeneity problem, this paper may shed new light on the relationship between schooling and juvenile delinquency by addressing the possible endogeneity. Furthermore, the research based on Japanese data can investigate if the results obtained in existing literature most of which based on US data can be generalised since the determinants of crime can be different in different countries.

The paper proceeds as follows: Section 2 presents theoretical and empirical literatures on the relationship between education and crime. In section 3 the econometric specification of equations are shown and several issues of estimation are discussed. Section 4 explains the earthquake and policies in focus. Section 5 provides definitions of variables and the data sources. Section 6 discusses the empirical findings, Section 7 checks robustness of the results and Section 8 concludes the paper.

2 Literature Review

2.1 Theoretical Studies

Ehrlich (1975) developed an economic model to explain the relationship between crime and education and empirically examined the model using US data. His explanation about an individual's participation in a criminal activity is based on cost-benefit analysis, i.e., an individual compares his expected income (utility) from engaging in legal and criminal activities and choose the activity which gives a higher income (utility).

Nearly a decade earlier, an economic model of criminal participation had been developed by other researchers such as Fleisher (1966) and Becker (1968). Ehrlich's (1975) contribution is to explicitly analyse the role of education in determining participation in the criminal activity within an economic model framework.

Ehrlich (1975) regards education as an efficiency parameter which affects returns from both legal and criminal activities as well as providing self-protection against apprehension or punishment. More education may increase returns from both legal and criminal activities as well as productivity of self-protection against apprehension or punishment. The effect of education on crime hence depends on the relative size of increase in production of returns from the legal activity and criminal activity and on the size of increase in production of self-protection. If education increases the productivity of legal activity more than that of criminal activity — ceteris paribus — more education would reduce an individual's incentive to engage in crime. A key element which determines the effect of education on crime is the degree of complementarity of education with inputs used to produce returns from legal and criminal activities.

The idea of Ehrlich (1975) is formalized by other researchers who add microfoundation to the intuitions provided above. Huang et al. (2004) developed an economic model to explain firstly the negative correlation between educational attainment and the crime rate and secondly the positive correlation between the unemployment rate and crime rate observed in some societies. Their model is a search model based on a two-stage process where at the first stage a household chooses a level of education through which it accumulates human capital and at the second stage the household chooses its career either from formal employment or crime.

Lochner (2004) developed a two-period life-cycle model of crime which investigates a dynamic interaction among crime, work and education choices. In his model, individuals with more human capitals commit less unskilled crime since they earn higher income in a formal labour market, and hence, their opportunity costs from forgone work is higher and their expected costs of committing crime associated with incarceration is higher. Lochner's (2004) mathematical explanation for this mechanism is summarised below.

Assumptions

Consider an individual who optimally chooses how much time to allocate in each period to investment in human capital, legal work, and crime to maximise his expected lifetime income. The individual is endowed with an initial skill level H_0, learning ability A, and criminal ability θ. He can choose to work, invest in human capital, and commit crime for the first T years of life. The individual can earn $w_t H_t + \varepsilon_t$ per unit of time spent working, where w_t represents the wage rate and ε_t is a mean zero iid shock. If he engages in crime, he may be incarcerated in the future with some probability. If incarcerated, the individual is provided a minimum level of consumption and cannot invest in human capital, work, or engage in crime until he is released. The investment in human capital produces future skills according to

$$H_{t+1} = H_t + f(I_t, H_t; A)$$

where H_t denotes a skill level at age t and I_t is time investment in a skill at age t. $f(\cdot)$ is increasing and concave. Individuals with higher learning ability receive higher returns on investments, i.e., $\frac{\partial^2 f}{\partial A \partial I} > 0$. The key assumption about learning here is that it is costly in terms of current income. In addition, there are direct costs of investment λ.

$N(k_t, H_t, \theta, \eta_t)$ is a net return of crime at period t where k_t denotes time spent committing crime and η_t is a mean zero iid shock to criminal returns. An individual who commits crime may be imprisoned at the beginning of the next period with probability $\prod(k_t)$ where $\prod(0) = 0$, $\prod(h) \leq 1$ and $\prod^{'}(k) > 0$. Note that total time each period is normalized to h so time spent working is equal to $h - I_t - k_t$. A convicted criminal must spend J years in prison and receives consumption c each year while he is in prison. While in prison, skills may depreciate, either from lack of use or through a stigma effect caused by prison, at the rate $\delta = [0, 1]$ per year. After a prison sentence ends, individuals are released, and again they choose to work, invest or engage in crime.

$V_t(H_t, \Sigma_t)$ is the expected value function for an individual who is not incarcerated at the beginning of period t conditional on his current state, where $\Sigma_t = (\varepsilon_t, \eta_t)$ is his current shock. $\Omega_t(H_t)$ is the expected value function for an individual who has just entered prison. These functions represent expected lifetime earnings at age t conditional on incarceration status, current human capital, and current shocks. An individual not in prison at the beginning of period $t \leq T$ decides how much time to allocate for work, investment in human capital and crime to maximise his expected lifetime income as described below:

$$V_t(H_t, \Omega_t) = \underset{I_t, k_t}{max}\{(w_t H_t + \varepsilon_t)(h_t - I_t - k_t) + N(k_t, H_t, \theta, \eta_t) - \lambda I_t - \prod(k_t) F$$

$$+ \beta[\prod(k_t)\Omega_{t+1}(H_{t+1}) + (1 - \prod(k_t))E(V_{t+1}(H_{t+1}, \Sigma_{t+1}))]\}$$

subject to

$$I_t, k_t \geq 0$$

$$0 \leq I_t, k_t \leq h, \; for\; any\, t$$

$$H_{t+1} = H_t + f(I_t, H_t; A)$$

where β is time discount factor. From ages $T + 1$ to $\overline{T}$, the final period of life, individuals are assumed to earn a certain income proportional to their human capital if they are not already incarcerated. Since old individuals do not commit crime, they will never face a new arrest during this period.

Relationship between Education and Crime

The optimality conditions for decision-making non-prisoners are as follows. The first order condition with respect to I_t:

$$w_t H_t + \varepsilon_t + \lambda = \beta[\prod(k_t)\Omega^{'}_{t+1}(H_{t+1}) + (1 - \prod(k_t))\frac{\partial E(V_{t+1}(H_{t+1}, \Sigma_{t+1}))}{\partial H_{t+1}}]\frac{\partial f}{\partial I_t}$$

which states that a marginal cost of time investment in skills is equal to a marginal value of time investment. As long as the marginal value of human capital for new prisoners is less than that for non-prisoners, $\Omega^{'}_t(H_t) < \frac{\partial E(V_t(H_t, \Sigma_t))}{\partial H_t}$, human capital offers a higher marginal payoff for those who are not in prison. Hence, criminals have less incentive to invest in human capital compared to non-criminals (reverse causality). Since the marginal value of human capital depends on future returns from work and crime, investment decisions will also depend on the subsequent work and crime decision:

$$\frac{\partial N(k_t, H_t, \theta, \eta_t)}{\partial k_t} = w_t H_t + \varepsilon_t + \prod^{'}(k_t) F$$

$$+\beta \prod^{'}(k_t)[E(V_{t+1}(H_{t+1}, \Sigma_{t+1}) - \Omega_{t+1}(H_{t+1})]$$

meaning that a marginal return from time investment in crime is equal to a marginal cost of time investment in crime. Assuming $dN/dH = 0$, a marginal return from crime increases less than an increase in a marginal cost of crime when a skill level H_t rises. As a result, conditional on ability and other permanent characteristics, an individual with high educational attainment faces high opportunity cost of crime and hence, on average, commits less crime.

Witte (1997) points out a different possible channel than an increased-opportunity-cost-channel through which education may affect the crime rate. She argues that education may change an individual's time preferences or provide information about "rules of the game", i.e., what is moral or legal. Further, she claims that education may provide information about costs and benefits of legal and criminal activities. Hence more education may affect an individual's crime decision by enabling the individual to judge what is moral or by enabling the individual to compute appropriately the cost of comitting crime. Witte (1997) proposes to analyse the above intuition in consumer demand model framework in which external information affects the parameters of the utility function and thus demand. Moreover, she points out a possible indirect effect of education on crime participation: more time spent in school may provide an individual with less crime prone friends, making the individual less crime prone.

2.2 Empirical Studies

Empirical studies which make inferences on the relationship between education and crime are not that numerous. First, Tauchen, et al. (1994) focus on males born in 1945 and residing in Philadelphia between their 10th and 18th birthdays and estimate effects of high school graduation and time spent in school on the probability of arrest without controlling for the endogeneity of schooling. The data is collected from school records, draft registration records, the Philadelphia Police Department, the Federal Bureau of Investigation (FBI), a compendium on city government finances, the Philadelphia Community Renewal Program, and interviews carried out in 1970-1971. Their results from probit model suggest that high school graduation status does not have a significant effect on the probability of arrest, whereas a proportion of time allocated to school during each year has negative and significant effect on the probability of arrest.

Second, Lochner (2004) uses cross-section data obtained from the National Longitudinal Survey of Youth (NLSY) and estimates the effect of schooling on crime without accounting for the endogeneity of schooling. Using a probit model in which a dummy variable indicating high school graduation status is used as the indicator of schooling, Lochner (2004) shows that high school graduation reduces participation in a criminal activity.

Third, Lochner and Moretti (2004) use the data obtained from the US census, the FBI crime report and the NLSY and conduct three different analyses two of which are micro level and one of which is aggregate level analysis. They account for the endogeneity of schooling using changes over time in the number of years of compulsory education that states mandate as an instrument for schooling. In the first analysis based on US census panel data, they show that schooling reduces the probability of incarceration, using a dummy variable indicating high school graduation status as an indicator of schooling. In the second analysis based on the NLSY cross-sectional data, they use self reported years of schooling and high school graduation status as indices of schooling and show that schooling reduces self reported participation in criminal activities. In the third analysis based on US census and FBI crime report aggregate panel data, they confirm the robustness of the results obtained in micro level analyses: higher average education or higher high school graduation rates in a state are associated with lower arrest rates. Lochner and Moretti (2004) is the only paper which accounts for the endogeneity of schooling.

3 Empirical Methodology

Estimation of the causal relationship between schooling and juvenile delinquency requires caution since the causal variable in focus, a schooling index, is potentially endogenous: that is, unobserved heterogeneity among individuals such as a variation in learning ability or criminal ability may affect schooling and crime participation decisions of the individuals. Moreover, the more an individual spends time committing crime, the less the individual may want to spend time for schooling due to the increased probability of being arrested before the benefits of schooling can be reaped.

This study examines the relationship between schooling and crime by regressing crime indicators on an index of schooling, controlling for various socio-economic indicators. Since micro data is not available, we use city level data to make a causal inference on effects of schooling on crime.[1] The following model is set up:

$$y_{it} = \alpha_i + x_{it}^{'}\beta + u_{it} \tag{1}$$

where i indicates the city and t denotes the year. y_{it} represents the crime indicator, α_i is the time invariant unobservables that vary over the city units i, and u_{it} is the disturbance term. β is $K \times 1$ vector, where K is the number of regressors and corresponds to coefficients to be estimated. x_{it} is $K \times 1$ vector and represents the index of schooling and other socioeconomic indicators. After controlling for these variables, the coefficient on the schooling index is expected to be negative.

An econometric issue in the estimation of model (1) is a possibility of endogeneity of the schooling index due to unobserved heterogeneity and simultaneity. A consistent estimator for β can be obtained by using the Within Group (WG) estimator to estimate model (1) if the only source of endogeneity is a correlation between the city specific time invariant unobservables and schooling index. In addition to this issue, if there is a correlation between the disturbances and schooling index, the Within Group Instrumental Variable (WG-IV) estimator should be used to estimate model (1) to get consistent estimator for β. The estimation of model (1) using the WG-IV estimator requires an instrument which gives an exogenous variation in schooling. In this paper, policies to support junior high school students, teachers and schools implemented after the Great Hanshin Awaji Earthquake will be employed as an instrument for schooling. Detailed explanations about the Great Hanshin Awaji Earthquake and the policies are as follows. [2]

[1] Fortunately, the work of Lochner and Moretti (2004) suggest that the results based on aggregate data and micro data are qualitatively similar.

[2] It is worth noting the work of Cipollone et al. (2007) since their estimation technique is relevant to this paper. Cipollone et al. (2007) focus on the effect of schooling on youth mortality, using a policy intervention triggered by a quake occurred in the Southern Italy in 1980 as an instrument for schooling. The policy allowed certain cohorts of males in specific towns affected by the quake not to be engaged in military service. Their work is based on aggregate level cross sectional analysis and the result suggests that an increase in high school graduation rates significantly reduces mortality

4 The Earthquake and Policy Intervention

4.1 The Great Hanshin-Awaji Earthquake

The Great Hanshin-Awaji Earthquake occured on 17th January 1995 with its epicentre at the southern part of Hyogo prefecture. The location of Hyogo prefecture and the epicentre are shown in Figure 1. The quake recorded magnitude of 7.2 and the highest seismic intensity of 7 on Japanese intensity scale at the southern part of Hyogo. The quake caused the largest loss in post-war Japan: 6,433 people were killed, 43,792 people were injured, and more than 300 thousands people evacuated. More than 398 thousands houses, factories, shops were destroyed or burnt down and infrastructures such as transportation system and port facilities were seriously damaged.

School facilities were also damaged: 3,883 schools including secondary schools were damaged. Many schools in the disaster area were forced to suspend classes after the quake and many students lost their houses and guardians. As a result, various measures to support students, teachers and schools were adopted.

4.2 Policies

4.2.1 Policy for Students

As one of various measures adopted to support students, the Government provided textbooks and school supplies with junior high school students in need of an emergency help due to the quake. The provision took place from January to March 1995 and 62,034 students received textbooks and school supplies.[3] The support was provided to junior high school students who could not go to schools due to a loss of textbooks and school supplies. Those who lived in cities and towns where the Government defined to be damaged according to the official evaluation performed shortly after the quake were eligible to obtain the support. These are 10 cities and 10 towns in Hyogo prefecture and 5 cities in Osaka prefecture.[4] The Government applied the Disaster Relief Act (DRA) to these cities and towns (hereafter, cities for brevity) and targeted these cities for providing various support. Figure 2 shows the location of cities to which the DRA was applied.

4.2.2 Policy for Teachers

The Government also adopted a measure to support schoolteachers and their families. Schoolteachers and their families in Hyogo and Osaka prefectures were exempted from paying medical bills (hospital charges and treatment fees) from January to May 1995 and as a result, 10,802 exemptions took place.[5] Schoolteachers and their families were eligible for the exemption if they lived in cities where the DRA was applied and (i) their residences are destroyed by more than 20% in terms of total area or (ii) their family members are dead or injured.

---

rates of youth.

[3] The figure includes the number of junior high school students and pupils.

[4] Specifically, these are Kobe-shi, Amagasaki-shi, Nishinomiya-shi, Ashiya-shi, Itami-shi, Takaraduka-shi, Kawanishi-shi, Akashi-shi, Miki-shi, Sumoto-shi, Tsuna-cho, Awaji-cho, Kitaawaji-cho, Kazunomiya-cho, Higashiura-cho, Goshiki-cho, Nishi-awaji-cho, Mihara-cho, Midori-cho, Minamiawaji-cho, Toyonaka-shi, Osaka-shi, Ikeda-shi, Suita-shi and Minoo-shi.

[5] 1,877 exemptions were recorded for schoolteachers working for private school, whereas the number of exemptions for schoolteachers working for public schools is unknown. Hence we compute the number of exemption for public school teachers assuming that the number of exemptions per school is identical in private schools and in public schools.

4.2.3 Policy for Schools

Schools in the disaster area were also supported by the Government. Junior high schools in the disaster area were subsidised to construct temporary classrooms and to reconstruct school builings. The construction of temporary classrooms started from February 1995 and as a result, 394 classrooms and 255 school buildings for junior high schools were (re)constructed. Schools were eligible to receive the subsidies if the total repair cost of school buildings or construction cost of temporary classrooms per school were above the following thresholds:

Threshold repair/construction costs (in a thousand Japanese yen)

Type of schools	National	Prefectural	Civic	Private
Junior High Shool	≥ 600	≥ 800	≥ 400	$\geq 1{,}500$

Not only threshold repair/construction costs but also subsidy rates vary with types of schools: the subsidy rate is 100% for national schools, 67% for prefectural and civic schools and 50% for private schools. Although this policy does not have a clearly defined target area, schools which received subsidies mostly concentrate on the DRA applied cities.

4.3 Empirical Issues

Due to the quake, many schools in the disaster area were forced to suspend classes, a large number of schoolteachers are injured, and residences of students are damaged. The policy intervention is, therefore, expected to have raised school attainment (or to have prevented the school attainment from dropping) in the disaster area by providing junior high school students with a normal studying environment even in the aftermath of the quake.

However, there is an empirical issue to identify the effect of treatments on schooling due to the nature of the policies. Specifically, since individuals and schools which received the treatments were suffered by the quake at the same time, the effect of treatments could be canceled out by the damage of the quake and we may not observe any effect of treatments on schooling. Therefore, it is essential to disentangle the effect of treatment and the effect of damage on schooling.

To identify the treatment effect, we use the fact that not all cities damaged by the quake were treated. Figure 3 shows a distinction of cities in Hyogo, Osaka and Kyoto prefectures according to their treatment and damage status. The shaded area indicates cities damaged by the quake and treated (DT group), whereas the lined area indicates cities damaged by the quake but not treated (DC group). The rest of the areas in Figure 3 where neither shaded nor lined are cities neither damaged nor treated (NDT group). We compare DT group and DC group to identify the effect of treatments on schooling.

A city is classified as treated if the Government defined to be damaged (the DRA applied cities) and a city is classified as damaged if the city recorded positive level of damage measured by percentage of death, injured and destroyed houses. Cities below the 25th percentile of each damage indicator are regarded as not damaged since these cities recorded fairly minor damage and hence it is unlikely that schooling in these cities are deteriorated by the quake damage.[6]

[6]The estimation result is not sensitive to this choice of cut-off value, i.e., the result is qualitatively same to the case in which cities with positive levels of damage is used as a definition of damaged cities.

4.4 Potential Effects of the Policy Intervention

In this section, timing of potential effects of policy intervention on schooling is discussed. Before moving on to detailed discussion about the timing, note that Japanese junior high school is a three-year-course and Japanese academic year starts from 1st of April and ends on 31st of March, e.g., academic year 1995 starts from 1st of April 1995 and ends on 31st of March 1996.

First, the policy for the junior high school students was implemented in January 1995 and continued until March 1995. Therefore, the policy is expected to stimulate demand for schooling in academic year 1994. Second, exemption from medical bills for schoolteachers and their families started in January 1995 and continued until May 1995. The exemption is, hence, expected to encourage supply of schooling in academic year 1994 and 1995. Third, construction fees of temporary classrooms and reconstruction fees of school buildings were subsidised from February 1995 until the (re)construction is completed. The subsidy is hence expected to stimulate supply of schooling in academic year 1994 and 1995.

Therefore, the policy intervention is expected to raise the high school participation rate in academic year 1995, 1996 and/or 1997 by encouraging the demand and supply of schooling in junior high schools in academic year 1994 and 1995. Evolution of the high school participation rate is shown in Figure 4. Three lines correspond to the average high school participation rates over time of DT, DC and NDT group, respectively. Figure 4 indicates that the average high school participation rate of each group shows upward trend. In 1995 and 1996, the average high school participation rate is evolving smoothly in DC group, whereas that of DT group is relatively increasing. Figure 4 supports a possibility that the high school participation rate increased in DT group but not in DC group. Robustness of the result will be checked after controlling for trend and other variables affecting the evolution of schooling in each group.

5 Description of Data and Variables

Crime data for 146 police offices in Hyogo, Osaka and Kyoto prefectures from 1990 to 2000 are taken from the Statistical Crime Report 1990 to 2000 (*Hanzai Tokeisyo*). The juvenile crime rate is defined as the number of juvenile crimes per 1,000 youth.[7] The crime level is measured in two ways: the recorded number of crimes and the number of individuals arrested. Both measures are recorded by crime category and by police office but not recorded by age. Fortunately, the number of arrested youth by crime category for each prefecture is available. Hence, the juvenile crime level for each crime category and police office is computed under the assumption that the proportion of arrested youth to the total number of arrests is uniform across police offices within a prefecture.

Crime levels are divided into four categories: total crime, violence, total theft and non-trespass theft. Total crime includes all acts in violation of the criminal law. Violence consists of murder, robbery, rape and arson. Total theft consists of trespass theft, vehicle theft and non-trespass theft. Trespass theft is the unlawful taking of property from the possession of another by entering a structure to commit a theft; vehicle theft is the unlawful taking of vehicles, e.g., motor vehicle thefts, from the possession of another; and non-trespass theft is any theft except trespass theft and vehicle theft such as shoplifting, purse-snatching and luggage lifting.

[7]Crimes committed by individuals aged between 14 to 19 are classified as juvenile crimes in the Japanese criminal law.

Data for covariates is available for 217 cities in Hyogo, Osaka and Kyoto prefectures from 1990 to 2000. Since the data for crime and for covariates have different units of observations, these data are matched in the following manner.

Case 1: One police office covers several cities

In Case 1, the data for covariates in levels are summed up, and then, the data for covariates in rates are computed subsequently. For example, assume that police office A covers cities B and C. When we compute the unemployment rate defined by the share of the unemployed to total labour force, first, the number of the unemployed and the number of total labour force, respectively, in cities B and C are summed up. Then, the total number of the unemployed in cities B and C is divided by the total number of labour force in cities B and C. Subsequently, the unemployment rate obtained is matched with crime data of police office A.

Case 2: One city is covered by several police offices

In Case 2, crime data in levels are summed up, and then, crime rates are computed subsequently. For example, assume that city A is covered by police offices B and C. When we compute the crime rate defined by the number of crimes per 1,000 youth, first, the number of crimes reported to police offices B and C are summed up. Then, the total number of crimes reported in police offices B and C is divided by the total number of youth in city A (and multiplied by 1,000), and subsequently, the crime rate obtained is matched with covariates data of city A.

As a result of matching, the number of cities and police offices become 112. Therefore, the sample consists of observations for 112 cities and police offices in Hyogo, Osaka and Kyoto prefectures from 1990 to 2000.

Schooling data are obtained from the School Basic Survey Report 1990 to 2000 (*Gakko Kihon Chosa Hokokusyo*). The survey covers every school in Japan including junior high schools and asks the number of classes, students, teachers, budget of school and careers of graduates etc. The survey has been conducted every year since 1948 and is a good source of community level information about schooling. The high school participation rate, defined as a share of junior high school graduates who go on to high school, is taken from the survey report.

All other data for covariates are obtained from a report of a survey, the System of Social and Demographic Statistics of Japan (*Syakai Jinko Tokei Taikei*). The survey covers all municipalities in Japan and collects information about socio-economic and demographic conditions of municipalities. The survey has been conducted since 1976 and provides 1,500 social and demographic variables by municipality.

As covariates, the wealth level, job opportunity and welfare generosity are included. The wealth level is measured by income per capita and job opportunity is measured by the unemployment rate. Poverty and less job opportunity would reduce expected income from legal activities and hence is expected to raise crimes. The welfare generosity is measured by logged welfare expenditure per capita. Welfare would reduce an incentive to commit a crime by providing an individual with an extra income.

Table 1 provides summary statistics for before and after the quake by group. Table 1 indicates that both the arrest rate and crime rate are lower after the quake in all groups. Income per capita, the unemployment rate and welfare expenditure per capita are similar in DT and DC groups, whereas those in NDT group are lower than in DT and DC groups.

6 Result

First, we begin by estimating model (1) using the GLS estimator and WG estimator to deal with a possible correlation between the unobserved heterogeneity, α_i, and the schooling index. Table 2 shows the estimates of regressions in which the number of arrested youth per 1,000 youth (hereafter, the arrest rate for brevity) is used as the dependent variable, whereas Table 3 shows estimates of regressions in which the number of recorded juvenile crimes per 1,000 youth (hereafter, the crime rate for brevity) is used as the dependent variable. The crime rate captures occurrences of crimes and hence it would reflect individual's criminal behaviour, whereas the arrest rate would reflect not only individual's criminal behaviour but also police's behaviour as not all criminals are arrested. The estimates provided in odd number columns are the WG estimates and those provided in even number columns are the GLS estimates.

The estimates in Table 2 indicates that the high school participation rate has a negative impact on the arrest rate. All estimates are statistically significant except the WG estimate in column 7. The estimates in Table 3 shows that schooling has a negative impact on the crime rate as well, although the estimates in total crime and total theft regressions are insignificant. To test if the unobserved heterogeneity is correlated with the schooling index, we conduct a hausman test under the null hypothesis that α_i and the schooling index are not correlated. The hausman test compares coefficient estimates in columns 1 and 2, 3 and 4, 5 and 6, and 7 and 8, respectively, of Table 2 and 3. The test rejects the null hypothesis at 5% level in all cases, suggesting that there is no reason to believe that α_i and the schooling index are not correlated. Hence, it would be important to account for the endogeneity due to the unobserved heterogeneity when the relationship between crime and schooling is estimated.

As yet a possible correlation between the disturbances, u_{it}, and the schooling index have not accounted for. If the disturbances and schooling index are correlated, the WG estimator is inconsistent. To account for the possible correlation between the disturbances and schooling index, the WG-IV estimator is used to estimate model (1). We begin by estimating the first stage equation of Two Stage Least Squares (TSLS) using the policy intervention after the quake as an instrument for schooling. Table 4 shows coefficient estimates of the first stage equation of the TSLS in which four variables are used as instruments: a dummy which takes 1 if a city belongs to DT group and year is 1995, and 0 otherwise (treatment dummy 1995); a dummy which takes 1 if a city belongs to DT group and year is 1996, and 0 otherwise (treatment dummy 1996); a dummy which takes 1 if a city belongs to NDT group and year is 1995, and 0 otherwise (damage dummy 1995); a dummy which takes 1 if a city belongs to NDT group and year is 1996, and 0 otherwise (damage dummy 1996). The omitted dummy is that for DC group. The treatment dummy and damage dummy, respectively, capture the effect of treatment and quake damage, respectively, on schooling, and hence, coefficients on both dummies are expected to be positive.

Table 4 shows that the coefficient estimate on treatment dummy 1995 is statistically not different from zero, whereas the estimate on treatment dummy 1996 is positive and significant, supporting a possibility that the treatments raised the high school participation rate in treated cities. The estimates on damage dummies 1995 and 1996, respectively, are both positive and significant, suggesting that the quake damage had a negative effect on the high school participation rate. The p value for the F test under the null hypothesis that the coefficients are jointly equal to zero is zero, implying that the instruments and other covariates have explanatory power to predict the high school participation rate.

Next, the second stage equation of TSLS is estimated. Table 5 shows coefficient estimates of regressions in which the arrest rate and crime rate, respectively, are employed as dependent variables. The estimates in Table 5 indicates that the high school participation rate significantly reduces both the arrest rate and crime rate of all crime categories. For instance, the estimate in column 6 of Table 5 indicates that 10% increase in the high school participation rate reduces the number of recorded juvenile total theft per 1,000 youth by 297 on average, ceteris paribus. This positive impact supports the theoretic prediction of Lochner (2004) suggesting that higher educational attainment reduces the participation in criminal activities through higher opportunity costs of committing crime or the intuitive explanation of Witte (1997) suggesting a possibility that higher educational attainment discourages criminal behaviours by increasing knowledge about what is moral/legal or increasing knowledge about costs of illegal activities.

To test if the disturbances, u_{it}, are correlated with the schooling index, a hausman test is performed under the null hypothesis that the disturbances and the schooling index are not correlated. The test compares coefficient estimates obtained by the WG-IV estimator and WG estimator. The null hypothesis is rejected in four cases among eight, suggesting that there is no reason to consider that u_{it} and the schooling index are not correlated in the regressions where the null is rejected. It would be hence important to account for the endogeneity due to simultaneity as well as the endogeneity due to unobserved heterogeneity.

Several results unrelated to schooling are also worth mentioning. First, income per capita shows a negative effect on both the arrest rate and crime rate, except the arrest rates of total crime and total theft. The result supports a possibility that more income discourages youth to commit crimes through higher opportunity costs of committing crimes. Second, the unemployment rate is positively related to both the crime rates and arrest rates, except the arrest rate of non-trespass theft. The result may imply that a higher unemployment rate encourages the youth to participate in crime by lowering expected income from legal market activities. Welfare expenditure shows a negative effect on both the arrest rate and crime rate, though only the effects on the crime rate of total theft and non-trespass theft are significant. An interpretation is that the welfare reduces an incentive to commit a theft by providing an individual with an extra income from a non-criminal activity.

7 Robustness Check

In this section, we examine robustness of the estimation results to various assumptions implicitly made so far. First, by comparing DT group and DC group to identify the treatment effect, we are assuming that DT and DC group are identical except treatment status after controlling for covariates. In other words, an implicit assumption is that the degree of damage affected only the assignment of treatment but it affected nothing else (recall that cities in DT group recorded more serious damage than cities in DC group).

To address a possibility that the degree of damage not only affects assignment of treatment but also schooling, police behaviour and other factors influencing the opportunity cost of committing crime, we retain in the sample only the least damaged cities in DT group, the most damaged cities in DC group, and cities in NDT group. A city in DT group is classified as the least damaged if the city recorded below the 50th percentile of each damage indicator in DT group. Damage is measured by percentage of death, injured and destroyed houses as before. A city in DC group is classified as the most damaged if the city recorded above the 50th percentile of each damage indicator in DC group. This rule selects 20 cities in DT group located farthest away from the epicentre and 15 cities in DC group, 7 of which share the boarder with cities in DT group. Since these cities are geographically close and suffered by the similar degree of damage, they must have been exposed to similar changes in schooling, police resources and other factors influencing the opportunity cost of committing crime.

Table 6 shows the estimation result of the first stage equation of TSLS using this restricted sample. The result confirms robustness of the result obtained in Section 6: the coefficient estimate on treatment dummy 1995 is statistically not different from zero, whereas the estimate on treatment dummy 1996 is positive and significant, supporting a possibility that the treatments raised the high school participation rate. The estimates on damage dummies 1995 and 1996, respectively, are both positive and significant, suggesting that the quake damage reduced the high school participation rate. The p value for the F test under the null hypothesis that the coefficients are jointly equal to zero is zero, implying that the instruments and other covariates have explanatory power to predict the high school participation rate. The estimation result of the second stage equation of TSLS is provided in Table 7. The result is similar to the one provided in Section 6 in terms of signs and significance, though there is a variation in magnitudes of coefficient estimates. For instance, the estimate in column 6 of Table 7 indicates that 10% increase in the high school participation rate reduces the number of recorded juvenile total theft per 1,000 youth by 126 on average, ceteris paribus.

Second, using damage dummies 1995 and 1996 as instruments are valid only if the quake damage do not have a direct effect on crime. If the quake damage had a direct effect on crime, damage dummies do not satisfy the exclusion restriction. A zero covariance assumption of a damage dummy and crime indicator may fail if, for example, the quake had increased looting behaviour in the aftermath or if the police anticipated that crime may increase in the aftermath and put more resources to the disaster area. To check if the quake had a direct effect on crime, we control for damage status in the second stage equation of TSLS.

Table 8 shows WG-IV estimates of regressions in which the crime rate and the arrest rate, respectively, are used as the dependent variables. In Table 8, damage dummies show negative and significant effect on the arrest rates, whereas the damage has an insignificant effect on the crime rate. After controlling for a damage status, the effect of schooling on the arrest rate remains negative but becomes insignificant, whereas the effect on the crime rate remains robust, i.e., negative and significant. A possible interpretation is that the police had allocated more resources to the disaster area, which results in an increase in the arrest rate in the disaster area, whereas the quake damage did not have a significant effect on youth's criminal behaviour (incentive), which is reflected in an insignificant effect of the quake damage on the crime rate. Therefore, the effect of schooling on the crime rate which captures youth's criminal behaviour is robust, while its effect on the arrest rate is not as robust as the crime rate since the arrest rate captures not only individual's criminal behaviour but also police's behaviour.

8 Conclusions

In this paper, we explored the relationship between schooling and juvenile delinquency. To identify the effect of schooling, we use an exogenous variation in schooling caused by a policy intervention in specific cities after the Great Hanshin-Awaji Earthquake. Using the WG-IV estimator, an endogeneity problem of schooling caused by simultaneity and unobserved heterogeneity is addressed. Our results indicate that higher school attainment reduces crime committed by youth regardless of crime categories. This negative impact of schooling on crimes supports a theoretical prediction of Lochner (2004) who suggests that lower educational attainment raises the participation in crimes through lower opportunity costs of committing crimes or the intuitive explanation of Witte (1997) who points out a possibility that lower educational attainment encourages criminal activities through a lack of knowledge about what is moral/legal or knowledge about costs of involving in criminal activities.

The results imply that policies which encourage individual investments in schooling may raise not only the individual benefits but also the public benefits which have received less attention by policy makers. Aside from classical policies that strengthen formal deterrence to crime such as an increasing severity of sentences, policies that encourage schooling may be an effective tool to reduce juvenile delinquency.

References

[1] Becker, G. S. (1968). "Crime and Punishment: An Economic Approach." Journal of Political Economy 76(2): 169 - 217.

[2] Cipollone, P., D. Radicchia, et al. (2007). The Effect of Education on Youth Mortality, unpublished manuscript.

[3] Ehrlich, I. (1975). On the Relation Between Education and Crime. Education, Income, and Human Behavior. F. T. Juster. New York, McGraw-Hill Book Co.

[4] Fleisher, B. M. (1966). The Economics of Delinquency. Chicago, Quadrangle Books.

[5] Gottfredson, D. C. (1985). "Youth Employment, Crime, and Schooling: A Longitudinal Study of a National Sample." Developmental Psychology 21(3): 419-432. Grogger, J. "Market Wages and Youth Crime." Huang, C. C., D.

[6] Headquarters of Hyogo Prefectural Police (1990-2000). Statistical Crime Report (Hanzai Tokeisyo). Hyogoken Keisatsu Honbu. Hyogo.

[7] Laing, et al. (2004). "Crime and Poverty: A Search-Theoretic Approach." International Economic Review 45(3): 909 - 938.

[8] Lochner, L. (2004). Education, Work, and Crime: A Human Capital Approach, National Bureau of Economic Research, NBER Working Papers: 10478.

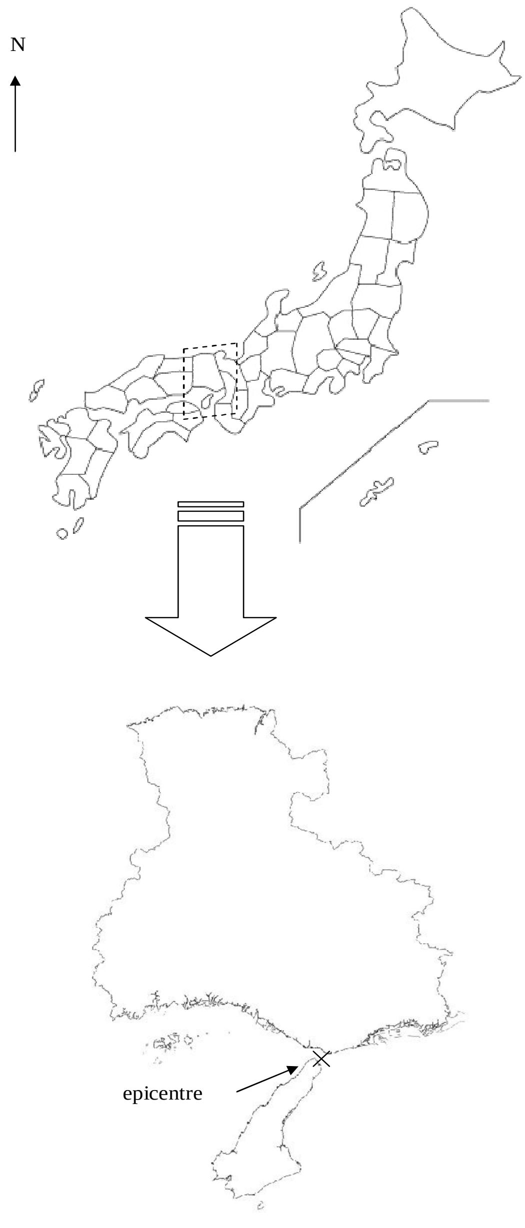
[9] Lochner, L. and E. Moretti (2004). "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." American Economic Review 94(1): 155-189.

[10] Ministry of Internal Affaires and Communications (2007). System of Social and Demographic Statistics of Japan (Syakai Jinko Tokei Taikei). Fundamental data of municipalities; Kinki block.

[11] Ministry of Education, Culture, Sports, Science and Technology (1990-2000). School Basic Survey Report (Gakko Kihon Chosa Hokokusyo).

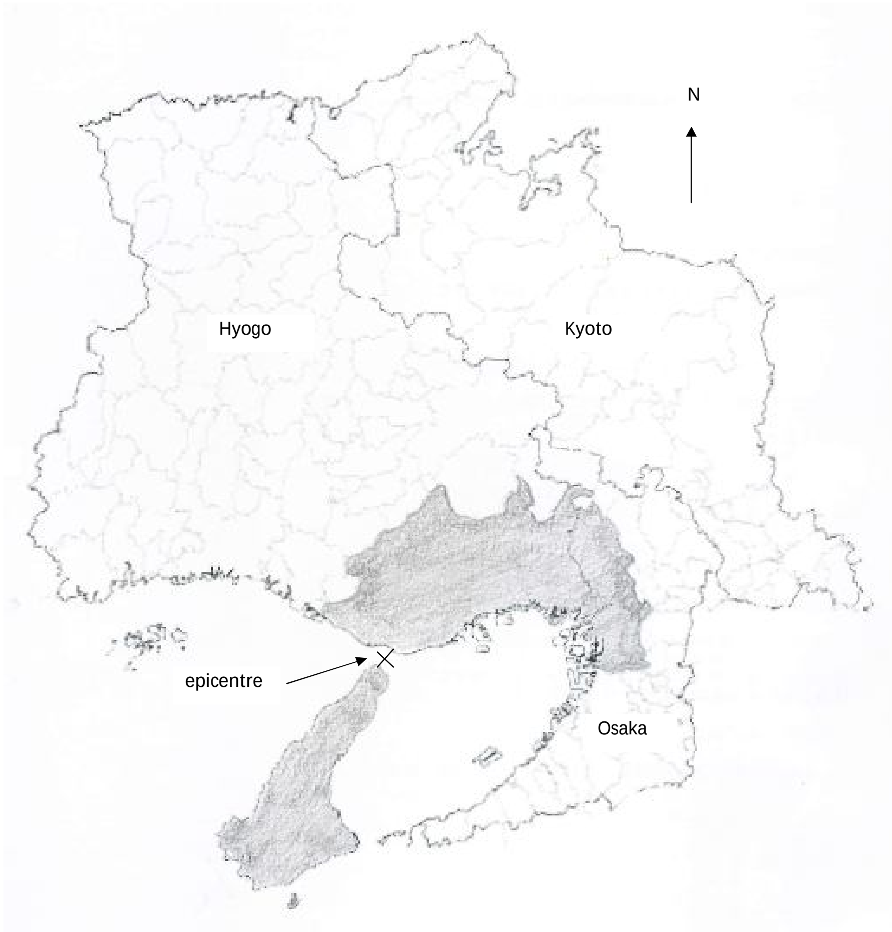
[12] Witte, A. D. (1997). Crime. The Social Benefits of Education. J. R. Behrman and N. Stacey, University of Michigan Press.

Figure 1: Epicentre



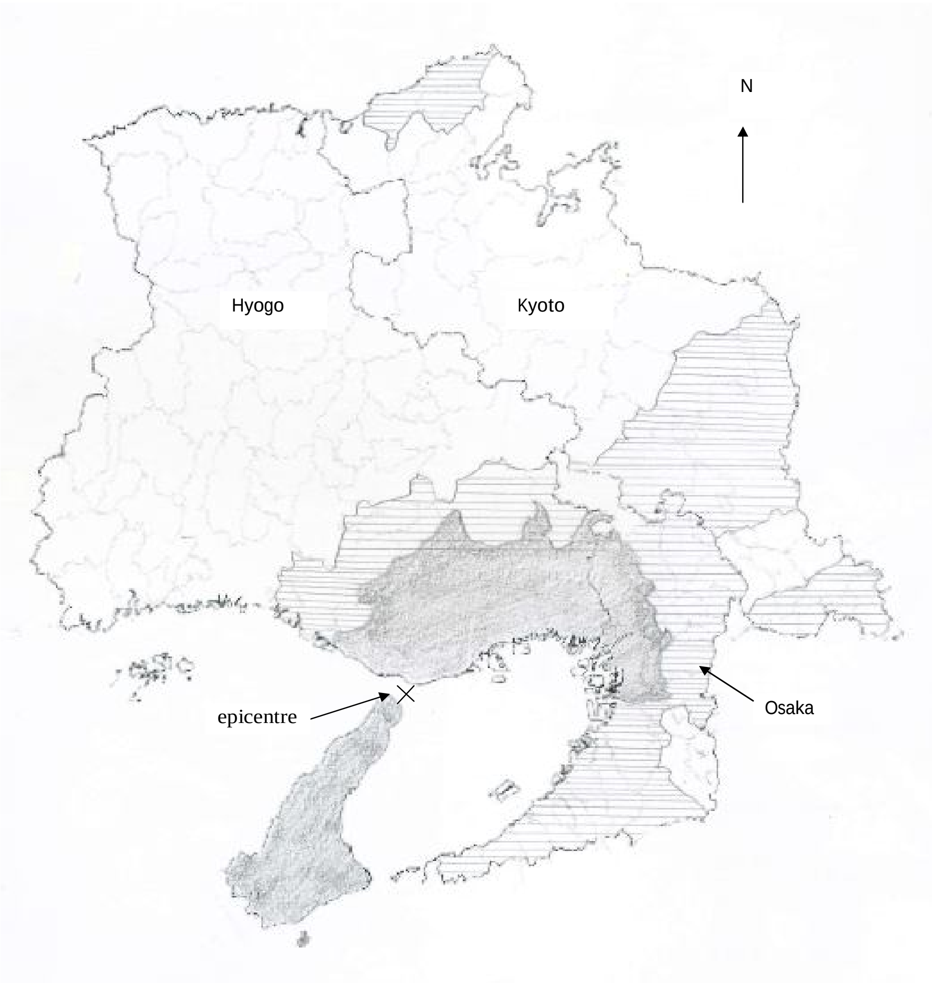
Note: the map at the top shows the location of Hyogo prefecture in Japan. The prefecture surrounded by the dotted square is Hyogo prefecture. The map at the bottom indicates the epicentre in Hyogo prefecture.

Figure 2: Disaster Relief Act Applied Cities



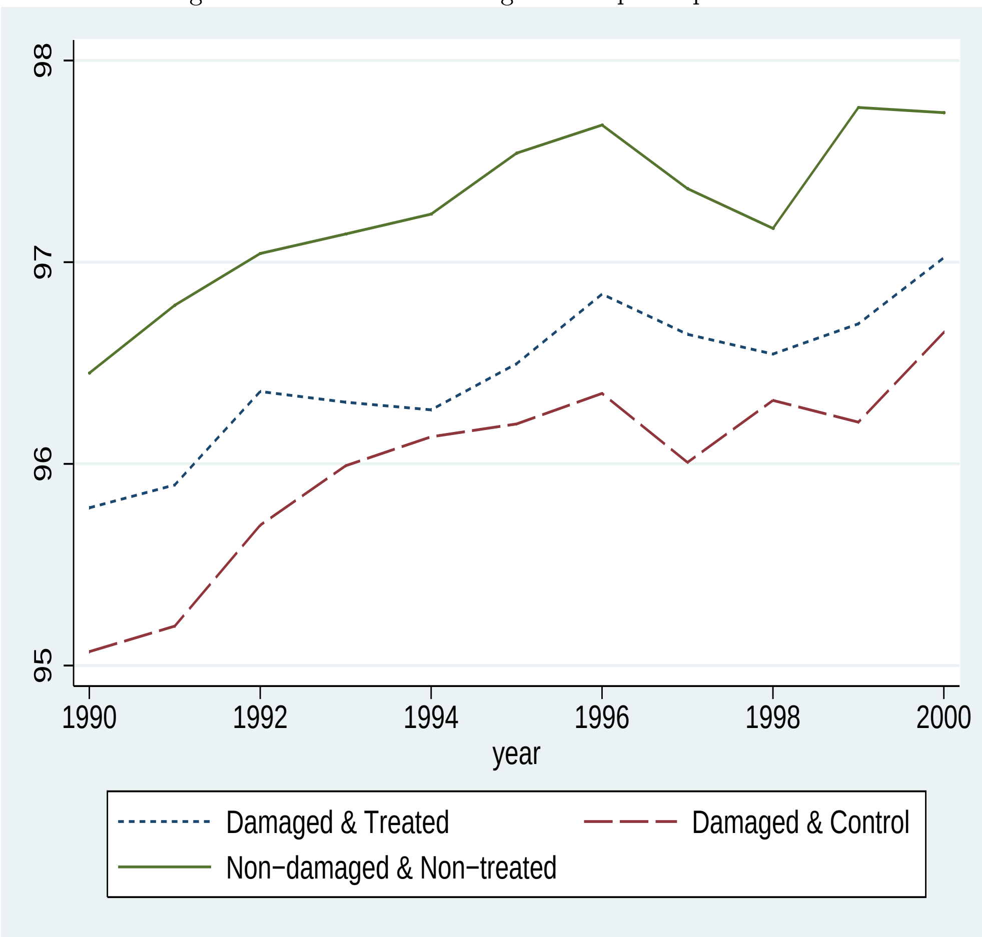
Note: the shaded area corresponds to 10 cities and 10 towns in Hyogo prefecture and 5 cities in Osaka prefecture where the Disaster Relief Act was applied.

Figure 3: Cities by damage and treatment status



Note: the shaded area corresponds to damaged & treated group, the lined area is damaged & control group, and the white area is non-damaged & non-treated group.

Figure 4: Evolution of the high school participation rate



Note: Three lines correspond to evolution of the high school participation rate by damage and treatment status.

Table 1: Summary Statistics

	Treatment Group			
	Post-quake		Pre-quake	
	Mean	Std. Dev.	Mean	Std. Dev.
Total Crime (arrest rate)	23.78	17.31	27.96	21.11
Total Theft (arrest rate)	14.26	10.34	15.40	12.06
Violent Crime (arrest rate)	0.15	0.12	0.43	0.50
Non-trespass Theft (arrest rate)	11.49	9.14	13.86	11.42
Total Crime (crime rate)	134.70	97.94	160.82	117.73
Total Theft (crime rate)	133.48	87.45	163.00	113.02
Violent Crime (crime rate)	0.19	0.15	0.50	0.54
Non-trespass Theft (crime rate)	118.12	79.74	145.78	102.68
Schooling	96.12	1.13	96.71	0.96
Income	7.34	0.19	7.34	0.16
Unemployment	3.41	3.30	4.50	4.28
Welfare	2.85	0.71	3.28	0.48
	Damaged & Control Group			
	Post-quake		Pre-quake	
	Mean	Std. Dev.	Mean	Std. Dev.
Total Crime (arrest rate)	15.71	3.43	19.10	4.88
Total Theft (arrest rate)	9.53	1.84	10.67	2.54
Violent Crime (arrest rate)	0.10	0.06	0.32	0.22
Non-trespass Theft (arrest rate)	8.07	1.84	9.60	2.32
Total Crime (crime rate)	97.82	23.37	119.80	36.83
Total Theft (crime rate)	100.05	23.07	126.12	39.40
Violent Crime (crime rate)	0.12	0.06	0.33	0.20
Non-trespass Theft (crime rate)	88.59	22.92	114.25	37.93
Schooling	95.62	1.61	96.29	1.68
Income	7.27	0.14	7.32	0.12
Unemployment	3.70	1.75	4.85	2.17
Welfare	2.61	0.37	2.95	0.26
	Non-damaged & Non-treated Group			
	Post-quake		Pre-quake	
	Mean	Std. Dev.	Mean	Std. Dev.
Total Crime (arrest rate)	12.89	5.11	15.13	6.86
Total Theft (arrest rate)	8.80	4.15	9.44	4.78
Violent Crime (arrest rate)	0.11	0.16	0.18	0.36
Non-trespass Theft (arrest rate)	6.85	3.87	8.34	4.84
Total Crime (crime rate)	65.06	25.93	71.10	34.61
Total Theft (crime rate)	66.63	26.13	74.91	37.39
Violent Crime (crime rate)	0.13	0.12	0.16	0.16
Non-trespass Theft (crime rate)	55.50	24.16	64.66	37.60
Schooling	96.93	1.34	97.54	1.19
Income	7.06	0.21	7.18	0.15
Unemployment	1.62	1.86	2.11	2.36
Welfare	2.24	0.88	2.70	0.65

Notes: pre-quake refers to years before 1995 and post-quake refers to year 1995 and after. Labels of variables indicate the following. Arrest Rate: the number of arrested youth per 1,000 youth; Crime Rate: the number of recorded juvenile crimes per 1,000 youth; Total Crime: violent plus property crime; Total Theft: trespass theft, non-trespass theft plus vehicle theft; Violent Crime: homicide, robbery, arson plus rape; Non-trespass Theft: theft without entering other individuals' properties such as shoplifting; Schooling: high school participation rates; Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita.

Table 2: Effects of Schooling on the Juvenile Arrest Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressand	total crime	total crime	total theft	total theft	violence	violence	non-trespass theft	non-trespass theft
Estimator	WG	GLS	WG	GLS	WG	GLS	WG	GLS
Schooling	-0.540**	-0.611***	-0.412**	-0.472***	-0.0289**	-0.0225***	-0.308	-0.396**
	(0.211)	(0.196)	(0.171)	(0.155)	(0.0116)	(0.00789)	(0.190)	(0.167)
Income	0.523	2.255	-1.580	0.103	-0.449***	-0.0661	-2.798	-0.127
	(3.195)	(2.542)	(2.469)	(1.870)	(0.137)	(0.0569)	(2.686)	(1.813)
Unemployment	0.918*	1.111**	-0.0968	0.136	0.0994***	0.0382***	-0.606*	-0.141
	(0.535)	(0.479)	(0.297)	(0.248)	(0.0204)	(0.00648)	(0.354)	(0.268)
Trend	0.468***	0.415***	0.215**	0.156*	0.0246***	0.0295***	0.428***	0.320***
	(0.132)	(0.117)	(0.0900)	(0.0816)	(0.00540)	(0.00351)	(0.103)	(0.0860)
Welfare	0.124	0.354	0.0327	0.244	-0.00628	0.00631	0.459*	0.748**
	(0.320)	(0.357)	(0.234)	(0.276)	(0.0109)	(0.0129)	(0.251)	(0.302)
Constant	62.31**	55.63**	62.20***	54.74***	5.806***	2.561***	58.70***	45.96**
	(26.06)	(24.28)	(20.66)	(19.20)	(1.248)	(0.786)	(21.05)	(19.18)
Observations	1190	1190	1190	1190	1190	1190	1190	1190
Num of cities	112	112	112	112	112	112	112	112
R-squared	0.170	0.164	0.031	0.075	0.206	0.201	0.102	0.048

Notes: Standard errors are clustered by city. *** denotes p<0.01, ** denotes p<0.05 and * denotes p<0.1. The dependent variable is the number of arrested youth per 1,000 youth. The labels of regressors indicate the following: Schooling: high school participation rates; Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita.

Table 3: Effects of Schooling on the Juvenile Crime Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressand	total crime	total crime	total theft	total theft	violence	violence	non-trespass theft	non-trespass theft
Estimator	WG	GLS	WG	GLS	WG	GLS	WG	GLS
Schooling	-0.349	-1.054	-0.366	-1.268	-0.0138*	-0.0165**	-0.562	-1.504*
	(0.869)	(0.776)	(1.049)	(0.908)	(0.00761)	(0.00741)	(0.871)	(0.783)
Income	-92.57***	-66.24***	-120.1***	-83.42***	-1.176***	-0.481***	-104.9***	-66.92***
	(16.03)	(11.82)	(18.15)	(12.17)	(0.267)	(0.116)	(16.07)	(10.86)
Unemployment	3.647	6.650***	4.641	8.112***	0.0831**	0.0534***	3.405	7.282***
	(2.862)	(2.291)	(3.501)	(2.623)	(0.0343)	(0.00835)	(3.021)	(2.249)
Trend	4.107***	3.331***	5.114***	4.150***	0.0374***	0.0345***	5.158***	4.109***
	(0.795)	(0.691)	(0.899)	(0.730)	(0.00990)	(0.00465)	(0.804)	(0.659)
Welfare	-2.193**	-0.194	-5.037***	-2.398*	-0.0137	0.0105	-5.514***	-2.704**
	(0.991)	(1.131)	(1.110)	(1.238)	(0.0100)	(0.0102)	(1.093)	(1.186)
Constant	790.9***	655.3***	992.8***	799.2***	9.662***	4.934***	893.7***	692.6***
	(98.42)	(82.14)	(113.2)	(93.37)	(2.201)	(1.201)	(113.0)	(89.54)
Observations	1190	1190	1190	1190	1190	1190	1190	1190
Num of cities	112	112	112	112	112	112	112	112
R-squared	0.292	0.106	0.353	0.105	0.368	0.223	0.353	0.131

Notes: Standard errors are clustered by city. *** denotes p<0.01, ** denotes p<0.05 and * denotes p<0.1. The dependent variable is the number of recorded juvenile crimes per 1,000 youth. The labels of regressors indicate the following: Schooling: high school participation rates; Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita.

Table 4: Effects of Treatment on Schooling

	(1)
Regressand	Schooling
Income	0.789**
	(0.357)
Unemployment	0.025
	(0.038)
Trend	0.099***
	(0.010)
Welfare	-0.059
	(0.064)
Treatment 95	0.001
	(0.135)
Treatment 96	0.365***
	(0.103)
Damage 95	0.349**
	(0.172)
Damage 96	0.35***
	(0.123)
Constant	90.28***
	(2.602)
Observations	1190
Num of cities	112
R-squared	0.25

Notes: *** denotes p<0.01, ** denotes p<0.05 and * denotes p<0.1. The dependent variable is high school participation rates. The labels of regressors indicate the following: Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita; Treatment 95: treatment dummy which takes 1 in 1995; Treatment 96: treatment dummy which takes 1 in 1996; Damage 1995: damage dummy which takes 1 in 1995; Damage 96: damage dummy which takes 1 in 1996.

Table 5: Effects of Schooling on the Juvenile Delinquency

	Arrest Rate				Crime Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressand	total crime	total theft	violence	non-trespass theft	total crime	total theft	violence	non-trespass theft
Schooling	-6.626***	-3.760***	-0.163**	-2.420**	-22.37***	-29.66***	-0.267***	-35.96***
	(1.838)	(1.171)	(0.0828)	(1.065)	(7.937)	(9.201)	(0.0811)	(9.855)
Income	5.163	0.973	-0.346**	-1.188	-75.78***	-97.72***	-0.983***	-77.93***
	(3.478)	(2.216)	(0.157)	(2.014)	(15.01)	(17.41)	(0.153)	(18.64)
Unemployment	1.136***	0.0233	0.104***	-0.530***	4.437***	5.692***	0.0922***	4.675**
	(0.345)	(0.220)	(0.0155)	(0.200)	(1.490)	(1.727)	(0.0152)	(1.850)
Trend	1.068***	0.545***	0.0378***	0.636***	6.281***	8.006***	0.0624***	8.651***
	(0.204)	(0.130)	(0.00920)	(0.118)	(0.882)	(1.023)	(0.00901)	(1.096)
Welfare	-0.0911	-0.0858	-0.0110	0.385	-2.972	-6.074**	-0.0226	-6.766**
	(0.585)	(0.373)	(0.0264)	(0.339)	(2.526)	(2.928)	(0.0258)	(3.137)
Constant	612.6***	364.9***	17.94**	249.7**	2782***	3641***	32.53***	4094***
	(167.8)	(106.9)	(7.558)	(97.18)	(724.5)	(839.9)	(7.399)	(899.6)
Observations	1190	1190	1190	1190	1190	1190	1190	1190
Num of cities	112	112	112	112	112	112	112	112

Notes: *** denotes $p<0.01$, ** denotes $p<0.05$ and * denotes $p<0.1$. The dependent variable is the number of arrested youth and recorded juvenile crimes, respectively, per 1,000 youth. The labels of regressors indicate the following: Schooling: high school participation rates; Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita.

Table 6: Effects of Treatment on Schooling with the Restricted Sample

Regressand	(1) Schooling
Income	0.299
	(0.484)
Unemployment	0.018
	(0.05)
Trend	0.107***
	(0.015)
Welfare	-0.137
	(0.087)
Treatment 95	-0.009
	(0.165)
Treatment 96	0.461***
	(0.165)
Damage 95	0.365**
	(0.177)
Damage 96	0.379***
	(0.126)
Constant	94.163***
	(3.511)
Observations	738
Num of cities	69
R-squared	0.245

Notes: *** denotes p<0.01, ** denotes p<0.05 and * denotes p<0.1. The dependent variable is high school participation rates. The labels of regressors indicate the following: Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita; Treatment 95: treatment dummy which takes 1 in 1995; Treatment 96: treatment dummy which takes 1 in 1996; Damage 1995: damage dummy which takes 1 in 1995; Damage 96: damage dummy which takes 1 in 1996.

Table 7: Effects of Treatment on Schooling with Restricted Sample

	Arrest Rate				Crime Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressand	total crime	total theft	violence	non-tresspass theft	total crime	total theft	violence	non-tresspass theft
Schooling	-5.782***	-2.808***	-0.204**	-2.147**	-8.410	-12.60*	-0.159*	-12.81*
	(1.778)	(1.087)	(0.0951)	(1.011)	(6.753)	(7.499)	(0.0835)	(7.187)
Income	5.767	0.215	-0.554***	0.472	-108.1***	-148.5***	-1.567***	-131.9***
	(3.946)	(2.411)	(0.211)	(2.244)	(14.99)	(16.64)	(0.185)	(15.95)
Unemployment	1.930***	0.216	0.0940***	-0.0800	5.803***	5.454***	0.0618***	4.770***
	(0.396)	(0.242)	(0.0212)	(0.225)	(1.503)	(1.669)	(0.0186)	(1.599)
Trend	0.804***	0.404***	0.0495***	0.489***	4.590***	6.561***	0.0657***	6.465***
	(0.215)	(0.131)	(0.0115)	(0.122)	(0.816)	(0.906)	(0.0101)	(0.868)
Welfare	-1.149	-0.631	-0.0306	-0.299	-3.366	-5.792*	-0.0155	-6.489**
	(0.723)	(0.442)	(0.0387)	(0.411)	(2.747)	(3.050)	(0.0340)	(2.923)
Constant	528.9***	280.4***	23.42***	212.8**	1671***	2366***	26.40***	2257***
	(166.0)	(101.4)	(8.873)	(94.39)	(630.2)	(699.8)	(7.797)	(670.7)
Observations	738	738	738	738	738	738	738	738
Num of cities	69	69	69	69	69	69	69	69

Notes: *** denotes p<0.01, ** denotes p<0.05 and * denotes p<0.1. The dependent variable is the number of arrested youth and recorded juvenile crimes, respectively, per 1,000 youth. The labels of regressors indicate the following: Schooling: high school participation rates; Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita.

Table 8: Effects of Schooling on the Juvenile Delinquency

	Arrest Rate				Crime Rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressand	total crime	total theft	violence	non-tresspass theft	total crime	total theft	violence	non-tresspass theft
Schooling	-2.406	-0.517	-0.0638	0.987	-17.13*	-29.97**	-0.379***	-39.65***
	(1.927)	(1.300)	(0.108)	(1.381)	(10.22)	(12.80)	(0.135)	(14.44)
Income	3.216	-0.520	-0.391**	-2.759	-78.25***	-97.58***	-0.929***	-76.18***
	(2.707)	(1.826)	(0.152)	(1.939)	(14.35)	(17.98)	(0.189)	(20.28)
Unemployment	1.034***	-0.0564	0.102***	-0.613***	4.322***	5.700***	0.0943***	4.751**
	(0.263)	(0.178)	(0.0148)	(0.189)	(1.396)	(1.749)	(0.0184)	(1.973)
Trend	0.635***	0.213	0.0277**	0.286*	5.740***	8.037***	0.0740***	9.033***
	(0.207)	(0.140)	(0.0116)	(0.148)	(1.097)	(1.374)	(0.0144)	(1.550)
Welfare	0.172	0.118	-0.00447	0.598*	-2.662	-6.093**	-0.0287	-6.978**
	(0.451)	(0.304)	(0.0254)	(0.323)	(2.389)	(2.993)	(0.0315)	(3.375)
Damage 95	-2.391*	-1.737*	-0.0287	-1.898*	-4.250	0.156	0.129	3.493
	(1.363)	(0.920)	(0.0767)	(0.976)	(7.227)	(9.055)	(0.0952)	(10.21)
Damage 96	-3.464***	-2.709***	-0.0947	-2.812***	-3.691	0.259	0.0606	2.360
	(1.084)	(0.731)	(0.0610)	(0.776)	(5.745)	(7.199)	(0.0757)	(8.119)
Constant	221.5	64.39	8.736	-66.05	2296**	3670***	42.89***	4436***
	(177.2)	(119.5)	(9.967)	(126.9)	(939.1)	(1177)	(12.37)	(1327)
Observations	1190	1190	1190	1190	1190	1190	1190	1190
Num of cities	112	112	112	112	112	112	112	112

Notes: *** denotes p<0.01, ** denotes p<0.05 and * denotes p<0.1. The dependent variable is the number of arrested youth and recorded juvenile crimes, respectively, per 1,000 youth. The labels of regressors indicate the following: Schooling: high school participation rates; Income: logged income per capita; Unemployment: unemployment rates; Trend: linear time trend; Welfare: logged welfare expenditure per capita; Damage 1995: damage dummy which takes 1 in 1995; Damage 96: damage dummy which takes 1 in 1996.