WPS4862

Policy Research Working Paper 4862

Natural Disasters and Human Capital Accumulation

Jesus Crespo Cuaresma

The World Bank
Sustainable Development Network Vice Presidency
Global Facility for Disaster Reduction and Recovery Unit
March 2009

Abstract

The author assesses empirically the relationship between natural disaster risk and investment in education. Although the results in the empirical literature tend to be inconclusive, using model averaging methods in the framework of cross-country and panel regressions, this paper finds an extremely robust negative partial correlation between secondary school enrollment and natural disaster risk. This result is exclusively driven by geological disasters. Natural disaster risk exposure is a robust determinant of differences in secondary school enrollment between countries, but not within countries, which implies that the effect can be interpreted as a long-run phenomenon.

This paper—a product of the Global Facility for Disaster Reduction and Recovery Unit, Sustainable Development Network Vice Presidency—is part of a larger effort in the department to disseminate the emerging findings of the forthcoming joint World Bank-UN Assessment of the Economics of Disaster Risk Reduction. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The task manager may be contacted at asanghi@worldbank.org, and the author of this paper at Jesus.Crespo-Cuaresma@uibk.ac.at.

Natural disasters and human capital accumulation[1]

Jesus Crespo Cuaresma[2]

Keywords: Natural disasters, human capital, education, school enrollment, Bayesian model averaging.

JEL Classifications: Q54, I20, E24, C11.

[1]This study was prepared as a background paper to the joint World Bank - UN Assessment on the Economics of Disaster Risk Reduction. Funding from the Global Facility for Disaster Reduction and Recovery is gratefully acknowledged. The author would like to thank Apurva Sanghi, the coordinator of the assessment, for many intellectually challenging discussions that helped shape the form of the paper. The paper has also profited from very helpful comments by José Miguel Albala-Bertrand, Jed Friedman, Samir KC, Norman Loayza, Reinhard Mechler, Paul Raschky, Gallina Vincelette and participants at the Brown Bag Lunch Seminar at the World Bank. All remaining errors should be exclusively attributed to the author.
[2]Department of Economics, University of Innsbruck; World Population Program, International Institute of Applied Systems Analysis (IIASA) and Austrian Institute for Economic Research (WIFO). Postal address: Universitätstrasse 15, 6020 Innsbruck, Austria. E-mail address: jesus.crespo-cuaresma@uibk.ac.at

1 Introduction

In this study we aim at quantifying the effect of natural disaster risk on investments in education by exploiting both cross-country and time differences in school enrollment rates. Given the large number of theories explaining differences in the rate of human capital accumulation across countries, we explicitly take into account model uncertainty by the use of model averaging techniques in order to extract the effect of catastrophic risk on school enrollment.

Traditionally, the empirical literature on the economic effects of natural disasters concentrated on the short-run effects of catastrophic events (see e.g. Dacy and Kunreuther, 1969). Albala-Bertrand (1993a, 1993b), Tol and Leek (1999), Rasmussen (2004) or Noy (2009) are some modern examples of research aimed at measuring the short and medium-run impacts of disasters. On the other hand, Skidmore and Toya (2002) and Crespo Cuaresma et al. (2008a) concentrate on long-run effects of disaster risk on the macroeconomy.[3] To the knowledge of the author, with the exception of some results in Skidmore and Toya (2002), there has been no fully-fledged empirical investigation on the effects of natural disasters on human capital accumulation hitherto. This is precisely the hole that this piece of research aims to fill.

From a theoretical point of view, the effect of natural disaster risk on educational investments is not unambiguous. Skidmore and Toya (2002) argue that to the extent that natural catastrophes reduce the expected return to physical capital, rational individuals would shift their investment towards human capital.[4] Although this argument highlights a possible channel of impact, this is just one of the possible effects of natural disasters on human capital. In principle, one could also argue that, in the framework of models of agents with finite lives, the potential effect that natural disaster risk has on mortality would lead to a lower level of educational investment in disaster-prone countries. Checchi and García-Peñalosa (2004) present a simple theoretical model which assesses the effect of production risk on education. In their model, aggregate production risk determines the average level of education, as well as its distribution. Checchi and García-Peñalosa (2004) show both theoretically and empirically that higher output volatility leads to lower educational attainment. Interpreting natural disaster risk as a component of aggregate production risk in the economy, countries which are more affected by disasters should also present lower levels of human capital accumulation, *ceteris paribus*.

The type of arguments put forward above stem from theoretical models and aim at unveiling the role of natural disaster risk as a determinant of cross-country differences. In this sense, these theoretical explanations refer to the long-run effects of natural disasters on educational investments. Short-run effects on human capital accumulation associated to the actual occurrence of the disaster could be extremely important as well. As a representative example,

[3] See Okuyama (2009) for a thorough review of the literature on the assessment and measurement of the economic effects of natural disasters.

[4] Skidmore (2001) studies investment decisions under catastrophic risk. Unfortunately, the empirical results in Skidmore (2001) are based on a dataset of very reduced size.

consider the 2005 earthquake in Pakistan. The Asian Development Bank and the World Bank (2005), when assessing the impact of the earthquake, estimate that 853 teachers and 18,095 students had lost their lives as a consequence of the disaster. They also report that the highest reconstruction costs in terms of material replacement requirements corresponded to schools from primary to higher secondary level. Over 7500 schools were affected by the earthquake and the estimated reconstruction costs for the education sector after the earthquake were the second highest by sector, after private housing.

Ultimately, the question of the human capital acummulation effects of natural disaster risk is an empirical one. The fact that we cannot rely on a single theoretical framework for the explanation of the link calls for an explicit assessment of the issue of model uncertainty when quantifying the effect of natural disasters on educational investments. In this contribution we use Bayesian Model Averaging (BMA, see Raftery, 1995, and Clyde and George, 2004, for general discussions on BMA and Fernández et al., 2001b, Sala-i-Martin et al., 2004, among others, for applications to the issue of the identification of robust determinants of economic growth) to obtain robust estimates of the effect of disaster risk on secondary school enrollment rates. The application of BMA ensures that our results are not specific to the choice of a particular model, and take into the account not only uncertainty of the estimates for a given model, but also uncertainty in the choice of a specification.

Our results indicate that geological disaster risk is a very robust variable for explaining differences in secondary school enrollment rates across countries. The effect is sizeable and well estimated. It implies that the school enrollment effect corresponding to the mean geologic disaster risk is around 1.65 percentage points in secondary school enrollment as compared to a country with zero disaster risk. The maximum disaster risk-driven effect in our dataset implies approximately a 20 percentage points decrease in secondary school enrollment.

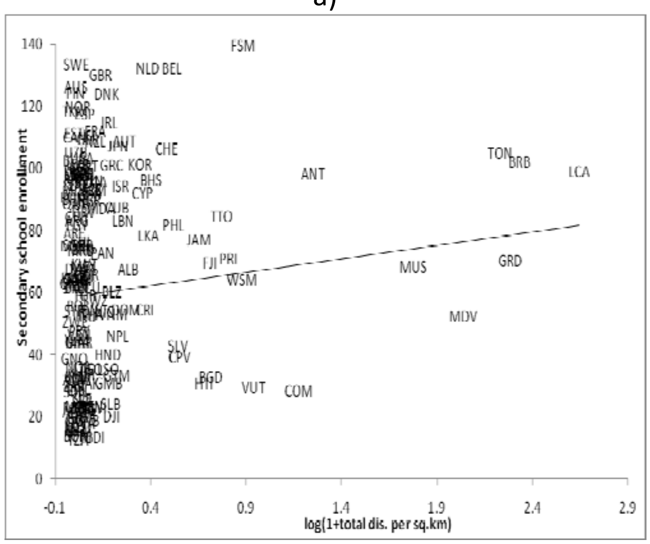
This contribution is structured as follows. Section 2 presents a first descriptive look at the empirical relationship between disaster risk and educational attainment. In section 3 we perform different BMA exercises to assess the robust effect of natural disaster risk as a determinant of differences in school enrollment rates both between and within countries. We also evaluate potential differences in the nature of the effect in the short versus the long run, and assess the issue of subsample heterogeneity in the response of human capital accumulation to disaster risk. Section 4 concludes.

2 A first look at education and disasters

In this section we present a first exploratory analysis of the relationship between natural disasters and human capital accumulation. Figure 1 presents scatterplots of average secondary school enrollment in the period 1980-2000 against different measures of natural disaster risk evaluated in the same period for 170 countries. As in Skidmore and Toya (2002), we concentrate for this graph on a simple measure of natural disaster risk based on average disaster occurrence by squared kilometer, so that disaster risk is measured as

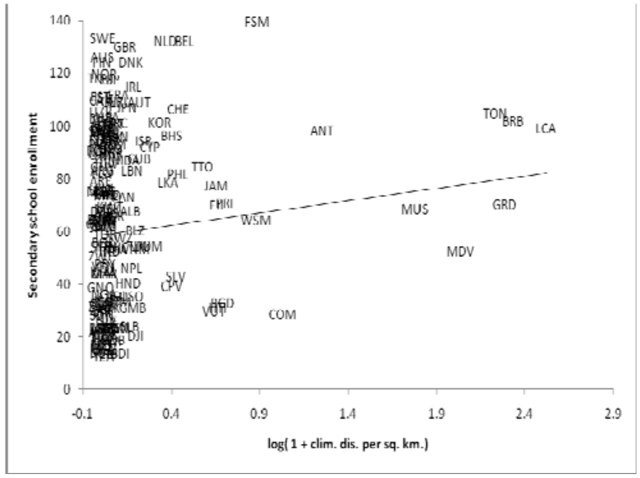
$d_i = \log[1+(\mathit{Number\ of\ disasters\ in\ country\ i/Area\ of\ country\ i}))]$.[5] We avoid using the existing data on quantified losses and received aid, since these type of measures are usually claimed to be plagued with endogeneity and other measurement problems. On the one hand, to the extent that disaster aid decisions are influenced by reported losses or affected people, governments would have an incentive to overreport these figures. On the other hand, the income level of a country (which is highly correlated with human capital accumulation) is a basic determinant of the effectiveness of natural disaster risk management. Since successful risk management mechanisms will reduce the negative macroeconomic effects of disasters, using estimated losses could spuriously lead to a negative correlation between disaster risk and education which is actually caused by the effect of education on the reduction of natural disaster loss. Skidmore and Toya (2007), for instance, show that higher levels of education reduce the losses from natural disasters. We therefore concentrate here on measures based on disaster occurrence frequency, which does not contain information on the magnitude of the disaster, but fulfills the necessary condition of exogeneity for the analysis.

a)



b)

[5]The source of disaster data is EM-DAT: The OFDA/CRED International Disaster Database (www.em-dat.net), Université Catholique de Louvain, Brussels (Belgium), where catastrophic events are reported which fulfill at least one of the following criteria: a) ten or more people reported killed, b) 100 people reported affected, c) a call for international assistance was issued or d) a state of emergency was declared.



c)

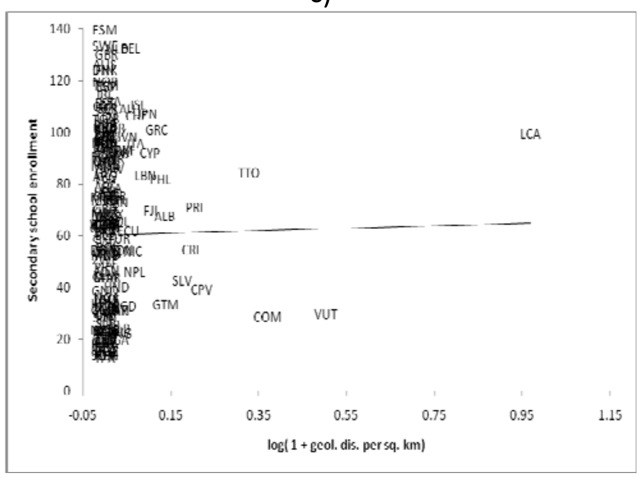


Figure 1: Natural disaster risk and secondary school enrollment: total disasters (a), geologic disasters (b) and climatic disasters (c)

In Figure 1 we show the raw correlation between disaster risk and education, without controlling for other factors that may affect school enrollment. We present scatterplots for all combined disasters, as well as for geologic and climatic disasters. We consider the following disasters as part of the group of climatic catastrophes: floods, cyclones, hurricanes, ice storms, snow storms, tornadoes, typhoons, storms, wild fire, drought, and cold wave, while the group of geologic disasters includes volcanic eruptions, natural explosions, avalanches, landslides, earthquakes and wave/surge. Since at this stage we exploit exclusively cross-country differences in these variables, we aim at interpreting the results as long-run effects of disasters on human capital accumulation. The results in Figure 1 indicate a weak positive relationship between disaster risk and the educational variable, which practically disappears for the subgroup of geologic disasters. Although the relationship is not statistically significant in any of the three cases reported in Figure 1, this first glimpse at the relationship we are interested in seems to lend some support to the conclusions in Skidmore and Toya (2002).

In order to extract the pure effect of disaster risk on educational investment, however, we need to control for a series of other variables which independently affect educational attainment differences across countries. Alternatively, a more precise picture of the relationship can be captured by concentrating on a more homogeneous group of geographical units. Although our study aims at a cross-country analysis, data at the regional level can deliver an interesting insight into the link between disaster risk and human capital accumulation. Figure 2 presents a scatterplot where the number of floods (as a rough measure of climatic disaster risk, log-transformed) is shown against the mean years of schooling of the population between 15 and 29 years of age for 75 districts in Nepal in the year 2001.[6] The relationship between educational attainment and disaster risk within this group of more homogeneous regions appears slightly negative, but not statistically significant.

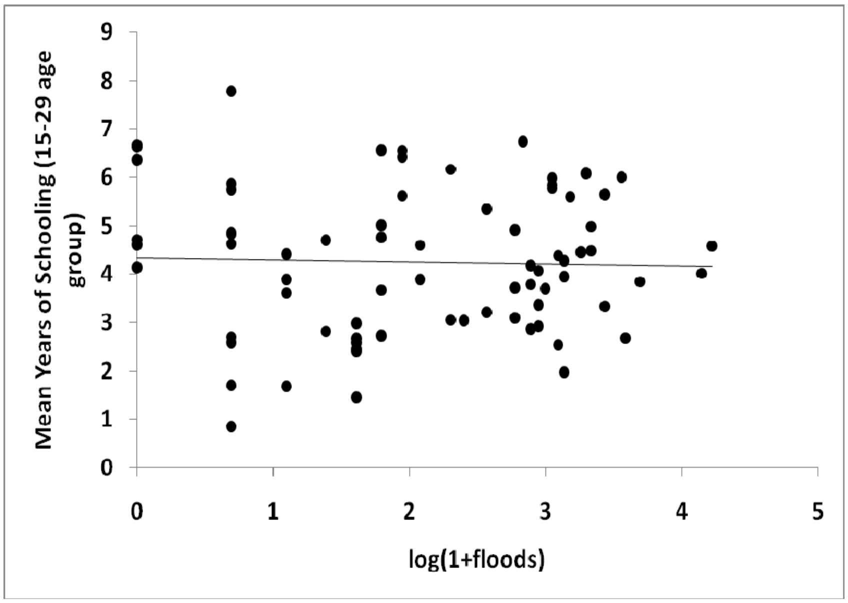


Figure 2: Floods versus mean years of schooling (age group 15-29) in Nepalese districts

Assessing the inmpact of natural disasters on education implies formulating a potentially large model, where a human capital accumulation measure is hypothesized to depend on a set of determinants and natural disaster risk. Obviously, the choice of extra controls to include in a model linking disaster risk to human capital accumulation depends on the theoretical setting assumed. When assessing empirically the issue of the determinants of human capital accumulation, we have many competing theories and effects that have been put forward in the literature to explain cross-country differences in educational attainment. So as not to make our empirical results depend on a specific theoretical (and thus econometric) specification , we investigate the issue of the robustness of disaster risk as a determinant of educational attainment in the framework of model uncertainty using BMA methods.

3 An empirical analysis of the effect of disaster risk on human capital accumulation

---

[6]The data stems from KC and Lutz (2009).

3.1 Model uncertainty

The estimation of the effect of catastrophic risk on human capital accumulation is carried out using linear econometric models of this type,

$$e_i = \alpha + \beta d_i + \sum_{j=1}^{k} \gamma_j x_j + \varepsilon_i, \tag{1}$$

where e_i is a proxy of educational attainment, d_i is the disaster risk variable, $X = (x_1 \ldots x_K)$ are other explanatory variables and ε is a zero-mean error term with variance equal to σ^2. In Skidmore and Toya (2002), for instance, the initial level of the educational variable as well as income per capita are the only variables in the **X** set. Since there are numerous variables affecting educational attainment, we aim at obtaining a measure which summarizes the effect of natural disaster risk on human capital accumulation after taking into account the degree of uncertainty which is embodied in the specification given by (1) when the nature of the variables in **X** which actually belong to the model and the size of the model itself are unknown.

Bayesian Model Averaging (BMA) presents a consistent framework for assessing the analysis put forward above.[7] Consider a set of $\overline{K}$ variables, $\mathbf{X}$ which are potentially related (linearly) to educational attainment in a cross-country regression framework, so that the stylized specification considered is given by (1) for $K \le \overline{K}$. In this situation there are $2^{\overline{K}}$ possible combinations of the regressors, each one defining a model M_k. The Bayesian approach to this problem implies considering model specification itself as a quantity to be estimated. In this sense, it follows immediatly that, by Bayes' theorem,

$$\mathrm{P}(M_k \mid \mathbf{Y}) = \frac{\mathrm{P}(\mathbf{Y} \mid M_k)\mathrm{P}(M_k)}{\sum_{m=1}^{2^{\overline{K}}} \mathrm{P}(\mathbf{Y} \mid M_m)\mathrm{P}(M_m)}, \tag{2}$$

which indicates that the posterior probability of model M_k is related to its marginal likelihood, $\mathrm{P}(\mathbf{Y} \mid M_k)$, and prior probability, $\mathrm{P}(M_k)$, as compared to the rest of models in the model space. Following Fernández et al. (2001a), we set an improper diffuse prior on α and σ, coupled with Zellner's (1986) g-prior on the $(\beta, \gamma_1, \ldots \gamma_k)$ parameter vector, which implies that

$$\mathrm{P}(\alpha, \beta, \gamma_j, \sigma \mid M_k) \propto \frac{1}{\sigma} N_{k+1}\left(0, \sigma^2\left((gX_{j'}X_j)^{-1}\right)\right), \tag{3}$$

where N_{k+1} is a multivariate normal distribution of dimension $k+1$ and X_j is a matrix whose columns are given by the independent regressors in model M_k. This setting implies that the

[7] Raftery (1995) and Clyde and George (2004) present general discussions of BMA in the setting of linear regressions.

Bayes factor (ration of marginal likelihoods) for two competing models, M_0 and M_1, is given by

$$B_{1,0} = \frac{P(\mathbf{Y}|M_1)}{P(\mathbf{Y}|M_0)} = \left(\frac{g}{g+1}\right)^{(k_1-k_0)/2} \left(\frac{1+g-R_1^2}{1+g-R_0^2}\right)^{-(N-1)/2}, \tag{4}$$

where N is the sample size, k_j is the dimension of model j and R_j^2 is the standard coefficient of determination for model j. Some particular values of g have been systematically used in the literature. For $g=1/N$ (the Unit Information Prior, UIP), the Bayesian Information Criterion (BIC) should be used as a basis of forming Bayes factors (see for example Kass and Wasserman, 1995, and Kass and Raftery, 1995), and thus BMA weights, while the Risk Inflation Criterion (Foster and George, 1994) sets $g=1/K^2$.[8]

Given $P(M_k|\mathbf{Y})$, we can build an estimate of the quantity of interest, say β, as a weighted average of all estimates of β where the weights are given by the posterior probability of each model from which the estimate was obtained,

$$E(\beta|\mathbf{Y}) = \sum_k E(\beta|M_k)P(M_k|\mathbf{Y}). \tag{5}$$

In a similar fashion, we can compute model averaged estimates of the posterior variance of β by computing the model averaged variance of the estimate, which in this setting summarizes information not only about the precision for a given model, but also across models.

3.2 The empirical setting

In our case, we will add as potential regressors in (1) a group of variables which have been put forward in the literature as important determinants of the differences in human capital accumulation across countries. Flug et al. (1998), for instance, assess the issue of the effect of macroeconomic volatility on investment in education, and present models which control for income inequality, credit market development, initial per capita income and initial educational levels. Some authors have also pointed out the importance of social and political institutions as factors affecting human capital accumulation (Stijns, 2006), while infrastructure variables can also thought of as having an effect on educational attainment. The group of potential regressors used in the BMA exercise are presented in Table 1. We concentrate on secondary school enrollment as the variable of interest. The focus on this measure is justified on the one hand by the fact that primary schooling is compulsory in most countries of our sample, and on the other hand because the most important results hitherto on the issue under study here (Skidmore and Toya, 2002) were obtained using gross secondary school enrollment as the human capital variable.

Table 1: Variables and definitions

[8] Fernández et al. (2001a) recommend using a benchmark prior on g based on the relative size of the group of potential regressors compared to sample size, so that $g=1/\max(\overline{K}^2, N)$.

Variable	Description	Source
e	Gross secondary school enrollment, average 1980-2000	WDI
e0	Initial gross secondary school enrollment, 1980	WDI
y0	Initial level of GDP per capita, 1980	PWT
gini	Gini index for income	WDI
life0	Life expectancy, 1980	WDI
vol	Volatility of GDP per capita growth	WDI
polity	Polity 2 indicator	PIV
pavroad	Percentage of paved roads	WDI
cred	Credit to private sector (% GDP)	WDI
war	Dummy variable for occurrence of war	--
d_t	Disaster risk, based on total disasters per thousand squared km	EMDAT
d_c	Disaster risk, based on climatic disasters per thousand squared km	EMDAT
d_g	Disaster risk, based on geologic disasters per thousand squared km	EMDAT

WDI: World Development Indicators (World Bank, 2006); PWT: Penn World Tables (Summer et al. 2006); PIV: Polity IV Database(Marshall and Jaggers, 1995); EMDAT: OFDA/CRED International Disaster Database (EMDAT, 2004).

As potential explanatory variables, we consider thus proxies of initial income and initial school enrollment ($y0$ and $e0$) in order to account, on the one hand, for wealth-induced human capital accumulation effects and for the observed persistence of human capital accumulation variables within countries, as well as for potential convergence in human capital across countries. In order to capture differences in income distribution across states we include also the country-wide Gini index, and the potential effect of macroeconomic instability is analyzed using the standard deviation of annual GDP growth rates as a proxy. We also control for differences in health by including life expectancy at birth in the initial period. Credit constraints are included into the model by using a proxy of financial depth: domestic credit provided to the private sector. The quality of (political) institutions is controlled for with the help of the Polity IV database, which offers a score variable (polity2) that combines two other score variables (Democ and Autoc) and quantifies the nature of the political system in a country. The Democ and Autoc scores include information on competitiveness and openness of executive recruitment, constraints on chief executive, regulation and competitiveness of participation. The polity2 measure is defined as Democ minus Autoc and ranges from +10 to-10 with -10 implying a strongly autocratic regime and +10 a strongly democratic regime. We also control for the existence of war in a given country during the period under study. The cross-country dataset contains data on all 80 countries for which all variables in Table 1 are available.[9]

Table 2: BMA results

	PIP/PM(PSD)	PIP/PM(PSD)	PIP/PM(PSD)
e0	*0.999/0.854(0.102)*	*0.999/0.850(0.102)*	*0.999/0.842(0.095)*
y0	*0.683/4.722(3.891)*	*0.694/4.834(3.889)*	*0.768/5.732(3.923)*
life0	*0.504/0.307(0.361)*	0.493/0.298(0.359)	*0.514/0.315(0.365)*
vol	0.242/-0.23(0.540)	0.239/-0.23(0.536)	0.249/-0.24(0.542)
polity	0.119/0.012(0.078)	0.118/0.011(0.077)	0.120/0.014(0.078)
pavroad	0.117/-0.00(0.016)	0.117/-0.00(0.016)	0.111/-0.00(0.015)
gini	0.273/-0.06(0.126)	0.272/-0.06(0.126)	0.256/-0.05(0.117)
cred	0.215/-0.01(0.044)	0.206/-0.01(0.043)	0.315/-0.03(0.059)
war	0.128/0.246(1.204)	0.126/0.232(1.182)	0.164/0.489(1.583)
d_t	0.287/-3.27(6.344)	-	-
d_c	-	0.173/-1.53(4.715)	-
d_g	-	-	*0.822/-51.9(31.16)*
g -prior	BIC	BIC	BIC
Prior model size	5	5	5
Obs.	80	80	80
Number of models	1024	1024	1024

PIP stands for ``Posterior inclusion probability", PM is ``Posterior Mean" and PSD is ``Posterior Standard Deviation". Figures in italics correspond to entries with PIP $>$ 0.5.

The results of the BMA exercise are presented in Table 2. We report the posterior inclusion probability of each variable (PIP), which is computed as the sum of the posterior probability of the models including that variable, together with the mean of the posterior distribution of the parameter attached to the variable (PM) and its standard deviation (PSD). The PIP can be interpreted as the probability that a given variable belongs to the true model, and we will classify explanatory variables as robust if after observing the data the probability that the variable belongs to the model increases with respect to the prior inclusion probability of the variable. For the BMA results in Table 2, we imposed a diffuse prior over the model space, so

[9] The countries included in the cross-sectional sample are (in abbreviation form): AUS, AUT, BDI, BEL, BFA, BGD, BOL, BRA, BWA, CAF, CAN, CHE, CHN, CIV, CMR, COL, CRI, DNK, DOM, DZA, ECU, EGY, ESP, FIN, FRA, GBR, GHA, GMB, GRC, GTM, HND, IDN, IND, IRL, IRN, ISR, ITA, JAM, JPN, KEN, KOR, LKA, LSO, MAR, MEX, MLI, MRT, MWI, MYS, NER, NGA, NIC, NLD, NOR, NPL, NZL, PAK, PAN, PER, PHL, PNG, PRT, PRY, RWA, SEN, SGP, SLE, SLV, SWE, SWZ, THA, TTO, TUN, TUR, UGA, URY, USA, VEN, ZMB and ZWE.

that $\mathrm{P}(M_f) = \frac{1}{2^K} \quad \forall f$, implying an average prior model size of $\overline{K}/2$ and a prior inclusion probability of 0.5 for all regressors. Each one of the columns in Table 2 presents the results for a group of regressors including a different natural disaster risk proxy in each one of the sets of covariates (all disasters, climatic disasters and geologic disasters, respectively). The initial educational attainment variable and the initial per capita income are highly robust in explaining secondary educational enrollment. The parameter attached to initial educational attainment is estimated very precisely and its posterior distribution has a mean below unity, implying (conditional) convergence in secondary school enrollment levels across countries. The initial income level variable also appears as a robust determinant of school enrollment, albeit with a very imprecise estimate of its effect. In two of the three settings presented in Table 2 life expectancy appears marginally as a robust determinant of educational attainment, with an effect which is positive but also unprecisely estimated. The effects of the other non-disaster variables appear neither robust in terms of PIP nor estimated with precision.

The results for the natural disaster risk variables shed a light on the nature of the channels between human capital accumulation and catastrophic risk. While the risk levels implied by using data on all disasters or climatic disasters do not appear robustly linked to school enrollment, if we consider exclusively geologic disasters the risk variable is robust and negatively linked to educational attainment. The effect is furthermore well estimated, with a ratio of PM to PSD of around 1.7. The results in the third set of columns imply that the decrease in school enrollment corresponding to the mean country in terms of geologic disaster risk is around 1.3 percentage points in secondary school enrollment as compared to a country with zero disaster risk. The maximum disaster risk-driven effect implies approximately a 15.9 percentage points decrease in secondary school enrollment.

3.3 Long-run or short-run effects?

The results presented indicate that natural disaster risk is a very robust variable for explaining differences in secondary school enrollment *across countries*. The natural question arises whether these are short-run or long-run effects. Does the occurrence of a disaster reduce schooling rates immediatly, so that the effect captured in the econometric analysis above is a direct consequence of the actual disaster realization? Or is the effect a longer-run, more structural type of change in the behaviour of economic agents when confronted with disaster risk? In order to answer this question, we redo the exercise above using two panels based on five and ten-year subperiods. This setting allows us to use country fixed effects and thus extract the potential within-country effect of disasters. The class of models we are considering in this exercise is thus

$$e_{it} = \alpha + \beta d_{it} + \sum_{j=1}^{k} \gamma_j x_{jt} + \varepsilon_{it}, \tag{6}$$

$$\varepsilon_{it} = \mu_i + \lambda_t + \nu_{it}, \tag{7}$$

where the error term ε can now be decomposed into a country-specific fixed effect (μ_i) which summarizes unobservable time-invariant country-specific characteristics, a fixed time effect common to all countries (λ_t) and the usual error term with constant variance (ν_{it}).

When implementing the BMA estimators we only consider models with the two-way fixed-effect structure in (7), so that the effects are to be interpreted as refering to the within-country dimension and therefore explain differences in school enrollment for a given country over time. The results of the exercise for the five-year and ten-year panels are presented in Table 3, which has the same structure as Table 2. The analysis in the panel setting does not include the variables *war*, which presents low variability in time, and *gini*, for which a single comparable observation was available for most countries in the sample.

Table 3: BMA results

5-year panel, within variation			
	PIP/PM(PSD)	PIP/PM(PSD)	PIP/PM(PSD)
$e0$	*0.974 / 0.415 (0.191)*	*0.974 / 0.415 (0.191)*	*0.975 / 0.414 (0.190)*
$y0$	0.100 / 0.435 (3.116)	0.100 / 0.434 (3.114)	0.100 / 0.428 (3.102)
$life0$	0.122 / -0.05 (0.313)	0.122 / -0.05 (0.313)	0.121 / -0.05 (0.310)
vol	*0.826 / -2.17 (1.743)*	*0.826 / -2.17 (1.743)*	*0.826 / -2.16 (1.740)*
$polity$	0.293 / -0.26 (0.603)	0.294 / -0.26 (0.604)	0.288 / -0.25 (0.594)
$pavroad$	0.072 / 0.001 (0.061)	0.072 / 0.001 (0.061)	0.072 / 0.001 (0.060)
$cred$	0.109 / 0.011 (0.067)	0.109 / 0.011 (0.067)	0.108 / 0.010 (0.066)
d_t	0.079 / 2.488 (33.57)	-	-
d_c	-	0.081 / 2.844 (34.12)	-
d_g	-	-	0.091 / -60.4 (493.9)
g -prior	BIC	BIC	BIC
Prior model size	4	4	4
Obs.	214	214	214
Number of models	256	256	256

10-year panel, within variation			
	PIP/PM(PSD)	PIP/PM(PSD)	PIP/PM(PSD)
$e0$	0.301 / 0.058 (0.136)	0.299 / 0.057 (0.136)	0.277 / 0.051 (0.130)
$y0$	*0.697 / 9.55 (9.934)*	*0.704 / 9.707 (9.974)*	*0.697 / 9.499 (9.886)*
$life0$	0.092 / 1.175 (15.53)	0.092 / 1.147 (15.48)	0.090 / 1.039 (15.24)
vol	0.236 / 0.221 (0.636)	0.234 / 0.218 (0.633)	0.238 / 0.223 (0.638)
$polity$	0.082 / -0.00 (0.117)	0.082 / -0.00 (0.117)	0.081 / 0.000 (0.117)
$pavroad$	0.085 / -0.00 (0.014)	0.085 / -0.00 (0.014)	0.084 / -0.00 (0.014)
$cred$	*0.930 / 0.156 (0.095)*	*0.931 / 0.157 (0.095)*	*0.940 / 0.161 (0.095)*
d_t	0.095 / 5.126 (54.70)	-	-
d_c	-	0.126 / 12.21 (69.85)	-
d_g	-	-	0.293 / -234. (560.4)
g -prior	BIC	BIC	BIC
Prior model size	4	4	4
Obs.	153	153	153
Number of models	256	256	256

PIP stands for ``Posterior inclusion probability", PM is ``Posterior Mean" and PSD is ``Posterior Standard Deviation". Figures in italics correspond to entries with PIP $>$ 0.5.

The results in Table 3 reveal that the robust negative effect of natural disaster risk on human capital accumulation found in the cross-country regression setting does not appear if we concentrate exclusively on within-country variation in school enrollment rates. Although the sign of the parameter attached to geological disasters remains negative, it is estimated with low precision and has an inclusion probability below 0.5. It is interesting to see that the inclusion probability of the disaster variables, and particularly for the geological disaster variable, increases when the horizon under consideration moves towards long-run comparisons. These results give us an interesting insight to the determinants of human capital accumulation in the short and medium run. The posterior inclusion probabilities of the variables for the five-year panel show that, apart from the natural persistence of human capital accumulation variables, income volatility plays the most important role as a determinant of secondary school enrollment rate changes over time. For the case of the ten-year panel, income developments and access to credit appear as the most robust variables. In particular the BMA estimate of $cred$ is very precisely estimated and implies that credit constraints play a privileged role in determining medium-run human capital accumulation dynamics. These results are in line and complement those found by Flug et al. (1998).

3.4 Parameter heterogeneity and interaction effects

An obvious question to ask in the framework of the analysis carried out in this paper is whether the effect of natural disaster risk on human capital accumulation depends on other country characteristics. The effects of natural disaster risk on several macroeconomic variables have been shown to be modulated by institutional and economic factors. Noy (2009) shows that the costs of disasters in terms of GDP depend on the strength of institutions in place in a given country, as well as on the level of income per capita. In a similar fashion, Crespo Cuaresma et al.

(2008a) also found that the potential positive effects of disasters on technology imports were only present for relatively developed countries, and did not exist in poor economies. The usual approach to the assesment of heterogeneity in elasticities is to include interaction terms, so that in this case the class of models considered (for the cross-country case) is given by

$$e_i = \alpha + \beta d_i + \eta d_i z_i + \sum_{j=1}^{k} \gamma_j x_j + \varepsilon_i, \tag{8}$$

where variable z (where in our case $z \in X$, although it need not be the case) is responsible for explaining differences in the elasticity of school enrollment to disaster risk.

There is some debate in the literature on how to treat interaction terms in the framework of variable selection and BMA. While some authors include the corresponding interaction as an extra linear covariate in the model, without setting any particular prior structure on models including the product of variables (see Masanjala and Papageorgiou, 2008), some other studies give a special treatment to models including interaction terms (see Chipman, 1996, for a general discussion and Crespo Cuaresma et al., 2008b and Crespo Cuaresma, 2009, for applications). The main problem with interpreting BMA results with interaction variables is that if the interaction term is considered a standard variable and we average over all possible combinations of variables we will be using some estimates based on models where the interaction is included but the interacted variables (the ``parent variables") are not part of the specification. These models do not allow us to interpret the interaction effect properly, since the absence of the parent variables in the specification implies that the interaction term may be actually capturing the direct effect of one or both of the parent variables. In this sense, if we aim at fulfilling what Chipman's (1996) *strong heredity principle*, we should only consider models where the interaction and the parent terms appear in the specification. For instance, in a more general setting, if there are K variables, where $K-1$ are standard variables and one is an interaction term (of variables in the former group), standard BMA would imply averaging over all 2^K possible combinations of these variables. On the other hand, the fulfillment of the strong heredity principle would require excluding from the model space those model specifications where the interaction without parent variables is included, which means that $2^{K-1}+2^{K-3}$ models would be actually evaluated.

Table 4: BMA results: Interaction terms

Cross-country		
	Standard BMA	Strong heredity fulfilled
	PIP/PM(PSD)	PIP/PM(PSD)
$d_t \times y0$	0.299 / -0.51 (1.676)	0.037 / -0.21 (1.674)
$d_c \times y0$	0.181 / -0.23 (1.302)	0.018 / -0.08 (1.167)
$d_g \times y0$	*0.525 / -4.15 (9.020)*	0.094 / -1.63 (10.94)
$d_t \times e0$	0.205 / -0.02 (0.087)	0.037 / 0.001 (0.048)
$d_c \times e0$	0.145 / -0.01 (0.070)	0.021 / 0.001 (0.040)

$d_g \times e0$	0.356 / -0.27 (0.550)	0.096 / 0.010 (0.314)
$d_t \times polity$	0.211 / -0.20 (0.543)	0.009 / -0.00 (0.135)
$d_c \times polity$	0.145 / -0.10 (0.424)	0.003 / -0.00 (0.077)
$d_g \times polity$	0.475 / -3.05 (3.883)	0.055 / -0.31 (1.512)
$d_t \times cred$	0.484 / -0.16 (0.202)	0.014 / -0.00 (0.039)
$d_c \times cred$	0.304 / -0.09 (0.173)	0.006 / -0.00 (0.029)
$d_g \times cred$	*0.751 / -1.34 (0.946)*	0.056 / -0.04 (0.340)
5-year panel		
	Standard BMA	Strong heredity fulfilled
	PIP/PM(PSD)	PIP/PM(PSD)
$d_t \times y0$	0.140 / 0.031 (0.147)	0.007 / 0.003 (0.053)
$d_c \times y0$	0.118 / 0.017 (0.109)	0.004 / 0.001 (0.036)
$d_g \times y0$	0.089 / 0.010 (0.240)	0.001 / 0.001 (0.070)
$d_t \times e0$	0.077 / 0.000 (0.000)	0.006 / -0.00 (0.000)
$d_c \times e0$	0.078 / 0.000 (0.001)	0.007 / -0.00 (0.000)
$d_g \times e0$	0.192 / 0.016 (0.052)	0.141 / 0.017 (0.052)
$d_t \times polity$	0.084 / 0.000 (0.009)	0.002 / 0.000 (0.005)
$d_c \times polity$	0.085 / 0.000 (0.009)	0.002 / 0.000 (0.005)
$d_g \times polity$	0.085 / -0.00 (0.067)	0.002 / 0.000 (0.030)
$d_t \times cred$	0.095 / 0.000 (0.001)	0.000 / 0.000 (0.000)
$d_c \times cred$	0.087 / 0.000 (0.001)	0.000 / 0.000 (0.000)
$d_g \times cred$	0.082 / 0.000 (0.006)	0.001 / 0.000 (0.001)
10-year panel		
	Standard BMA	Strong heredity fulfilled
	PIP/PM(PSD)	PIP/PM(PSD)
$d_t \times y0$	0.095 / -0.00 (0.028)	0.009 / -0.00 (0.026)
$d_c \times y0$	0.120 / 0.001 (0.027)	0.008 / -0.00 (0.024)
$d_g \times y0$	0.280 / -0.05 (0.302)	0.039 / -0.03 (0.320)
$d_t \times e0$	0.257 / -0.00 (0.005)	0.155 / -0.00 (0.005)
$d_c \times e0$	0.210 / -0.00 (0.005)	0.119 / -0.00 (0.005)
$d_g \times e0$	0.355 / -0.00 (0.012)	0.009 / -0.00 (0.003)
$d_t \times polity$	0.089 / -0.00 (0.005)	0.000 / -0.00 (0.000)
$d_c \times polity$	0.092 / -0.00 (0.006)	0.000 / -0.00 (0.000)

$d_g \times polity$	0.099 / 0.000 (0.028)	0.001 / -0.00 (0.001)
$d_t \times cred$	0.218 / -0.00 (0.003)	0.038 / -0.00 (0.001)
$d_c \times cred$	0.090 / 0.000 (0.002)	0.011 / -0.00 (0.000)
$d_g \times cred$	*0.962 / -0.02 (0.014)*	*0.802 / -0.02 (0.021)*

PIP stands for ``Posterior inclusion probability", PM is ``Posterior Mean" and PSD is ``Posterior Standard Deviation". Figures in italics correspond to entries with PIP $>$ 0.5. Columns labelled ``Strong heredity fulfilled" perform BMA using only models where the parent variables are included when the interaction is active.

We apply both approaches to our dataset in order to evaluate the existence of subsample heterogeneity in the effects of natural disaster risk on human capital. We evaluate different model spaces, each one containing potential interactions of the disaster variable with a) the level of income per capita, b) the initial level of school enrollment, c) the political regime and d) the degree of credit constraint, as measured by $cred$. Thus, BMA estimates are obtained for model spaces defined by the specifications in (8) with z given by each one of these variables. The results are presented in Table 4, where the PIP, PM and PSD for the interaction terms are presented.[10] Several interesting results emerge from the analysis. On the one hand, there is basically no evidence for robust heterogenous effects of natural disasters on education depending on the variables specified. If we concentrate on the results obtained by imposing the strong heredity principle, the only interaction with PIP over 0.5 corresponds to the combined effect of geological disasters and financial depth ($cred$) in the ten-year panel. The BMA estimate corresponding to this variable, although only marginally precise, indicates that, ceteris paribus, a country which is more (geological) disaster prone tends to react less in terms of school enrollment to the alleviation of credit constraints. A similar negative effect is found in the long-run cross-section for the standard BMA setting, although the effect is not at all robust if we consider only models where the interacion term appears exclusively together with the parent variables. In the cross-section setting the interaction of income and (geological) disaster risk is marginally robust, but the effect estimate is extremely imprecise and, just as in the case of credit constraints, it disappears if the strong heredity principle is imposed in the design of the BMA estimators.

3.5 Other robustness checks

We also performed other robustness checks to ensure that the results are not driven by the prior structure imposed on the BMA procedure. The results shown above are robust to changing the parameter prior from the Unit Information Prior to the Risk Inflation Criterion, as well as to the use of a hyperprior on model size as proposed by Ley and Steel (2009). For the cross-country setting, we also performed BMA on an alternative set of covariates which enlarges the group of explanatory variables in Table 1 by an extra variable that measures the percentage of mountainous terrain in the respective countries. With this variable, we control for geographical and topographical effects that may be correlated with the disaster risk variables but exert an independent effect on human capital investment (by, for instance, affecting the return of

[10] Complete results for all other variables are available from the author upon request. The results presented in section 3.2 are not affected by the inclusion of the interaction terms as extra variables.

infrastructure in terms of school enrollment). The BMA results concerning the importance and size of the effect of geological disaster risk were basically unchanged, while the mountainous terrain variable only achieved a posterior inclusion probability of 0.12. The posterior distribution of the parameter attached to the variable is, as expected, negative, but estimated with very low precision (the ratio of posterior mean to posterior standard deviation is -0.14).

We also explicitly assessed the issue of influential observations by estimating BMA parameters and inclusion probabilities based on subsamples. The results concerning the long-run effects of geological disaster risk on secondary school enrollment rates are robust to the following changes in the dataset:

- Trimming the observations corresponding to the disasters with the highest ratio of affected individuals per square kilometer (several percentiles of the distribution of affected people by area were tried as the cutting point, ranging from the 80th to the 95th, so as not to reduce the estimation sample dramatically)

- Trimming the observations corresponding to the poorest countries of the sample (thresholds based on observed income levels ranging up to the 30th percentile were tried)

- Trimming the observations corresponding to zero disasters, so as to ensure that the results are not driven exclusively by the differences between observations with zero versus positive disaster risk

- Excluding the 4 observations which are identified as outliers after inspection of the residuals of the specification which includes all ten potential variables.

4 Conclusions

The effects of natural disaster risk on human capital accumulation have received little attention in the academic literature hitherto. We offer here a first fully-fledged empirical study assessing the effects of natural disasters on secondary school enrollment rates. In order to avoid model-specific results, we use Bayesian model averaging techniques to assess the robustness and quantify the effects of natural disaster risk on human capital accumulation.

Our results give strong evidence of negative long-run effects of geological natural disaster risk on secondary school enrollment rates. The effects are only present in the short and medium run to the extent that natural disasters cause instability and increase output volatility. The long-run effects tend to be homogenous across countries and do not depend on income or the degree of human capital accumulation in the respective country. Furthermore, the political regime does not affect the disaster risk effect on human capital accumulation. The empirical results presented here are robust to numerous variations of the setting.

References

Albala-Bertrand, J. (1993a), Natural disaster situations and growth: a macroeconomic model for sudden disaster impacts, *World Development* 21, 1417-1434.

Albala-Bertrand, J. (1993b), *Political economy of large natural disasters,* Clarendon Press, Oxford.

Asian Development Bank and the World Bank (2005), Pakistan 2005 Earthquake Preliminary Damage and Needs Assessment. Islamabad, Pakistan, November 12, 2005.

Checchi D. and C. García-Peñalosa (2004), Risk and the distribution of human capital. *Economics Letters* 82, 53–61

Chipman, H. A. (1996), Bayesian Variable Selection with Related Predictors. *Canadian Journal of Statistics* 24, 17–36.

Clyde, M.A., and E.I. George (2004), Model Uncertainty, *Statistical Science* 19, 81–94.

Crespo Cuaresma, J., Hlouskova, J. and M. Obersteiner (2008a), Natural disasters as creative destruction: Evidence from developing countries, *Economic Inquiry* 46, 214-226.

Crespo Cuaresma, J., Doppelhofer, G. and M. Feldkircher (2008b), The determinants of economic growth in European regions. *Working Paper 2008-26, Faculty of Economics and Statistics, University of Innsbruck*.

Crespo Cuaresma, J. (2009), How different is Africa?. *Mimeo, Faculty of Economics and Statistics, University of Innsbruck*.

Dacy, D.C. and H.C. Kunreuther (1969), *The economics of natural disasters,* Free Press, New York.

Fernández, C., Ley, E. and M.F. Steel (2001a), Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics* 100, 381-427.

Fernández, C., Ley, E. and M.F. Steel (2001b), Model Uncertainty in Cross-Country Growth Regressions. *Journal of Applied Econometrics* 16, 563-576.

Flug, K., A. Spilimbergo and E. Wachtenheim (1998), Investment in education: do economic volatility and credit constraints matter? *Journal of Development Economics* 55, 465-481.

Foster, D.P. and E.I. George (1994), The Risk Inflation Criterion for Multiple Regression. *Annals of Statistics* 22, 1947-1975.

Kass, R.E. and A.E. Raftery (1995), Bayes factors. *Journal of the American Statistical Association* 90, 773–795.

Kass, R.E. and L. Wasserman (1995), A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. *Journal of the American Statistical Association*, 90, 928-934.

KC, S. and Lutz, W. (2009), Vulnerability to flood and landslide in Nepal. *Mimeo, International Institute for Applied Systems Analysis*.

Ley, E. and M.F. Steel (2009), On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression. *Journal of Applied Econometrics*, forthcoming.

Masanjala, W.H. and C. Papageorgiou (2008), Rough and Lonely Road to Prosperity: A Reexamination of the Sources of Growth in Africa Using Bayesian Model Averaging. *Journal of Applied Econometrics* 23, 671-682.

Marshall, M.G. and K. Jaggers(1995), Polity IV Project. Political Regime Characteristics and Transition, 1800-2004, Version 2005.

Noy, I. (2009), The macroeconomic consequences of disasters, *Journal of Development Economics*, forthcoming.

Okuyama, Y. (2009), Critical review of methodologies on disaster impact estimation, Background Paper to the joint World Bank - UN Assessment on the Economics of Disaster Risk Reduction.

Raftery A. E. (1995), Bayesian Model Selection for Social Research, *Sociological Methodology* 25, 111-163.

Rasmussen, T. N. (2004), Macroeconomic implications of natural disasters in the Caribbean, *IMF Working Paper WP/04/224*.

Sala-i-Martin, X. Doppelhofer, G. and R. Miller (2004), Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach, *American Economic Review* 94, 813-835.

Skidmore, M. (2001), Risk, natural disasters, and household saving in a life cycle model, *Japan and the World Economy* 13, 15-34.

Skidmore, M. and H. Toya (2002), Do natural disasters promote long-run growth?, *Economic Inquiry* 40, 664-687.

Skidmore, M. and H. Toya (2007), Economic development and the impacts of natural disasters, *Economic Letters* 94, 20-25.

Stijns, J-P. (2006), Natural resource abundance and human capital accumulation. *World Development* 34, 1060-1083.

Tol, R. and F. Leek (1999), Economic analysis of natural disasters, in *Climate, Change and Risk*, ed. T.E. Downing, A.J. Olsthoorn and R.S.T. Tol, London: Routledge.

Zellner, A. (1986), On assessing prior distributions and Bayesian regression analysis with gprior distributions, in *Bayesian Inference and Decision Techniques: Essays in Honour of Bruno de Finetti*, eds. P.K. Goel and A. Zellner, Amsterdam: NorthHolland, pp. 233–243.