An economic analysis of the impact of deforestation on flood risk in developing countries

Susana Ferreira and Ramesh Ghimire
Department of Agricultural and Applied Economics
University of Georgia
Athens, GA 30602, USA
sferreir@uga.edu

Aug 10, 2011

Abstract

We investigate whether deforestation increases flood frequency in developing countries. Using the data and sample of Bradshaw *et al.* (2007), but different statistical techniques, we obtain comparable results, namely natural forest cover is associated with a decrease in the count of floods. However this result is not robust to changes in the sample or to controlling for unobserved country heterogeneity. We further account for population, land use, and socioeconomic and governance characteristics, and find that income, corruption, population and urbanization indicators are significant determinants of flood frequency.

An economic analysis of the impact of deforestation on flood risk in developing countries

1. Introduction

Despite the conventional wisdom that forests reduce the frequency and magnitude of floods, the relationship between deforestation and large flood events remains controversial. In a recent paper, Bradshaw *et al.* (2007) analyze country-level data from 56 developing countries on flood characteristics, land cover, land cover change and geophysical controls, and conclude that deforestation amplifies flood frequency in developing countries. Reanalyzing the data used by Bradshaw *et al.* (2007), Van Dijk *et al.* (2009) conclude that deforestation "does not affect large flood events, although associated landscape changes can under some circumstances."[1] Van Dijk *et al.* (2009) argue that the frequency of large flood events is better explained by population, an omitted variable in Bradshaw *et al.*'s (2007) flood risk analysis, rather than by land cover change in the form of deforestation.

We reanalyze the number of floods between 1990 and 2000 using the same cross-section of developing countries and similar explanatory variables as in Bradshaw *et al.* (2007). Our results are consistent with theirs; we find that natural forest cover is negatively and significantly associated with flood frequency. We then extend their analysis of flood risk in three directions. First, we expand the sample to include all the developing countries for which we are able to gather data. Second, as suggested by Van Dijk *et al*. (2009), we account

[1] For a flood event to be considered "large" and recorded in the dataset used in both studies, it has to fulfill at least one of the following criteria: significant damage to structures or agriculture, long (decades) reported intervals since the last similar event, and/or fatalities (DFO 2011).

for population and landscape changes associated with urban population growth. Moreover, we incorporate other socioeconomic factors (income and corruption) in the analysis, and show that these factors play a significant role in driving flood frequency. Third, since the socioeconomic and flood data are available on a yearly basis between 1990 and 2000, we exploit the panel nature of the data and use country fixed effects to control for countries' unobserved heterogeneity. The negative impact of natural forest cover on flood risk is not robust to any of the three extensions.

More generally, our paper contributes to improve our understanding of the relationship between deforestation, socioeconomic factors and floods in developing countries. We believe that this is important for several reasons. First, floods are the most common natural disaster, accounting for 40 percent of the total natural disasters over the last 25 years (CRED/OFDA 2011), and the frequency of the most severe among the large flood events is increasing (Figure 1). Floods are also very costly. Each year they result in loss of lives, displacement, and large property and infrastructural damages. For example, in 2010, about 7,800 people died, and over 17 million people lost their home as a consequence of flooding (DFO 2011). However, there is little empirical work that addresses the flood-deforestation link for large floods at the country level, with most of the hydrological studies dealing with small scale or catchment-specific events.

Second, socioeconomic and institutional factors cannot be ignored when explaining flood frequency. We include them in our models. As noted by Van Dijk *et al*. (2009), controlling for population may prove fundamental. Humans have historically migrated to flood plains in search of water and fertile lands. Without adequate design of stream flow system, rapid

expansion of population in the flood plains may increase the frequency of floods, especially if natural vegetation is replaced by impervious surfaces. In addition, more population means a higher probability of a flood getting reported, with floods in sparsely populated area such as natural forests more likely to go unrecorded in media and official records. In our analysis we attempt to capture these effects by including both population and urban population growth indicators. Finally, flood frequency may depend on flood control infrastructure provided by the government, for example in the form of dams, dikes, and levees, which in turn depend on socioeconomic and institutional factors such as income and ability of the government to effectively provide them, for which we also account.

2. Forests, people and floods

According to the conventional wisdom, forests act as giant 'sponges', soaking up abundant rainfall and storing it before releasing it in regular amounts over an extended period (see e.g. CIFOR/FAO 2005 or Myers 1986), thus reducing the frequency and magnitude of floods. In addition, forests reduce rain-induced floods since they have higher evapotranspiration rates than alternative land uses, resulting in larger direct evaporation of rainfall intercepted by the canopy. Further, deforestation is associated with reduced soil infiltration and increased water run-off.

However, among hydrologists, the relation between deforestation and large flood events remains controversial (Reed 2002; Van Dijk *et al.* 2009). The consensus is that deforestation can have a role in flood formation at a small scale, or catchment-specific level (Tollan 2002), but that it does not play a significant role in the formation of large floods because these are the outcome of a number of factors such as geological composition, terrain slope, soil

permeability, porosity, crusting and prior wetness, and incident rainfall intensity and duration (Reed 2002). Recent reviews on this topic can be found in Bruijnzeel (2004), Calder (2007), Van Dijk & Keenan (2007).

In addition to geological factors and country's physical characteristics, land use and other human activities also influence the peak discharge of floods by modifying how rainfall is stored on and run off the land surface into streams. In undeveloped areas such as forests and grasslands, rainfall collects and is stored on vegetation, in the soil column, or in surface depressions. In contrast, urban areas, where much of the land surface is covered by roads and buildings, have less capacity to store rainfall. Construction of roads and buildings often involves removing vegetation, soil, and depressions from the land surface. The permeable soil is replaced by impervious surfaces such as roads, roofs, parking lots, and sidewalks that store little water, reduce infiltration of water into the ground, and accelerate runoff to ditches and streams. Likewise, dense networks of ditches and culverts in cities reduce the distance that runoff must travel overland or through subsurface flow paths to reach streams and rivers (Konrad 2005).

Infrastructures such as dams, dikes, and levees share the characteristics of public goods. The efficient provision of public good is, in turn, associated with both larger incomes and better institution (Deacon 2003). Recent studies use socioeconomic and institutional indicators to explain mortality resulting from natural disasters. They typically find that richer, more democratic and less corrupt countries experience fewer deaths after natural disasters because they are better able to agree on and invest in preemptive measures (better building codes, zoning restrictions), and better fund emergency and healthcare services required after the

disaster strikes (Kahn 2005). Compared with other natural disasters, however, floods offer more scope for policy intervention, not only for the mitigation of damage once the flood occurs, but for reducing the intensity of a flood or preventing it entirely.[2] Humans have actively managed rivers and their drainage basins for millennia. Income and institutions may determine the efficiency of flood-management-related actions which in turn affect the frequency and magnitude of floods (Congleton 2006; Ferreira 2010).

3. Data

We compiled data on large flood events, natural and non-natural forest cover, geophysical characteristics, population, income, corruption and land use change (urbanization) for 66 developing countries that experienced at least one large flood event, over the period 1990 to 2000.[3]

3.1 Flood Data

Flood related data are collected from the Dartmouth Flood Observatory (DFO 2011), a publically accessible global archive of large flood events. DFO uses different collection tools such as MODIS (Moderate Resolution Imaging Spectroradiometer, http://modis.gsfc.nasa.gov) and optical remote sensing, which provide frequent updates of water condition worldwide to detect and locate river flood events. In addition, DFO uses wide variety of news and governmental sources to complement them.

---

[2] Most earthquakes are caused by movement of the Earth's tectonic plates. Human activity can also produce earthquakes through the construction of large dams and buildings, drilling and injecting liquid into wells, coal mining and oil extraction, and nuclear tests, but these instances are very rare (Kisslinger 1976).
[3] Because of missing values for some observations, the actual sample sizes in the analyses are smaller.

Regarding the sample selection, as in Bradshaw *et al.* (2007), we only consider floods caused by heavy rain or brief torrential rain, excluding events caused by typhoons, cyclones, or other causes (e.g. dam failure) that originate independently of landscape characteristics. Water bodies are not confined to national boundaries, and neither are floods. In some cases, floods in a country could originate in neighboring countries. To facilitate the inter-country and within-country interpretation of our results we exclude multi-country floods from our analysis.

Our measure of flood risk or flood frequency is the total number of floods experienced by a country. On average, there was just over half a flood per country-year in the sample (Table 1), but the number varies considerably, with a standard deviation larger than twice the mean.

3.2 Forest cover

Since non-natural forest cover may not always affect water yields in the same manner as native vegetation (Bruijnzeel 2004), as Bradshaw *et al.* (2007), we introduce two distinct variables in the analysis: area covered by natural forest and area covered by non-natural forest. Both variables, measured in thousand hectares for years 1990 and 2000 come from FAO (2010). We converted them to square km.

3.3 Socioeconomic and institutional indicators

Population and urban population growth data come from the World Bank's World Development Indicators (WDI 2010). For income, we use GDP per capita, also from WDI (2010) measured in 2005 international dollars and adjusted to account for purchasing power

parity. Our corruption index comes from the International Country Risk Guide (PRS 2011). It rates from 0 to 6, with larger values denoting lower corruption.

3.4 Other controls

We control for a number of physical factors at the country level. We use the same controls as Bradshaw *et al.* (2007) (total area of the country, rainfall, slope, degraded landscape area, and soil moisture regime) although in some cases the data have been corrected and updated. For example, Bradshaw *et al*. (2007) use average annual precipitation data (in mm.) over 1950-1990 to capture the effect of rainfall on explaining flood frequency between 1990 and 2000. We instead use yearly data from 1990 to 2000 from the Tyndall Centre for Climate Change Research at the University of East Anglia (TCCRR 2011).

Regarding the other variables, land area in square km. is collected from WDI (2010). Average uphill slope of the country's surface area is from Nunn and Puga (2010). Degraded land, defined as area of each country devoted to urbanization, cropland and cropland/natural vegetation mosaic was obtained from the WRI (2011) . To obtain the soil moisture[4] regime we overlay a global soil moisture map, obtained from USDA (2011) with the world's physical boundary map in ArcGIS. The soil moisture regime of a country (aridic, xeric, ustic, udic, or perudic) corresponds to the largest area class of the country's soil that falls on the particular category. We further classify these 5 moistures regime into three regimes – arid,

[4] Soil moisture can have profound effects on frequency of floods. For example a relatively small amount of water accumulation from rainfall in an arid region can lead to flooding where as an equivalent amount of rainfall on per-humid soils may not result in any particularly noticeable accumulation of surface water (Beljaars *et al.* 1996; Cassardo *et al.* 2002).

sub-humid, and, per-humid following Bradshaw *et al*. (2007) and create dummies for each category.

Table 1 summarizes the descriptive statistics for all the variables and table 2 presents their simple correlation coefficients. Interestingly, simple correlation coefficients between the annual total number of floods and natural forest cover, positive, do not offer support to the hypothesis that deforestation increases flood frequency.

4. Estimation strategy

4.1 Revisiting Bradshaw et al.'s (2007) results

Our starting point is the estimation of the first 9 models considered by Bradshaw *et al.* (2007)[5] Table 3 summarizes the variables used in each of their models. Except in model 9 that does not include forest cover variables, in all the cases natural and non-natural forest cover variables are included separately. In addition, a variable capturing the natural forest loss (change in forest cover between 1990 and 2000) is considered. Models differ in how the forest cover variables are introduced; in levels, interacted (with slope and natural forest loss). They consider the second interaction plausible given that flood frequency may depend on the amount of forest loss relative to the initial state. The sample is also identical to that in Bradshaw *et al.*'s (2007) 56 developing countries, but due to missing observations the actual sample size is reduced to 39-51 countries depending on the specification.

[5] There are altogether 8 models or specification because model (7) and model (8) are exactly the same (see Bradshaw *et al.* 2007, Table 1, pp. 2385).

Despite flood data being available in a panel format, all the models in Bradshaw *et al*. (2007) are estimated with cross-sectional data. The dependent variable is the sum of all the floods occurred between 1990 and 2000. The independent variables are either time invariant (in the case of most of the physical characteristics) or averaged over the period (for rainfall and the forest cover variables).

4.2 Extensions

We extend the analysis of flood risk in three directions. First, we expand the sample to include China, and all the developing countries for which we are able to gather data. This increases the number of observations to 66, resulting in more degrees of freedom, in addition of more variation. We also test the robustness of the results to including all the countries, i.e. developing and developed. Appendix Table 1 contains the list of countries for each of the samples. Second, we incorporate the socioeconomic and institutional variables in the regressions. That is, we include variables for population density, urban population growth, income and corruption in the analysis. In the cross sectional regressions, averages for the period 1990-2000 are used. Third, we exploit the panel nature of the data.

Socioeconomic and flood data are available on a yearly basis between 1990 and 2000. Forest cover data are available only for the years 1990 and 2000, but we can interpolate the intermediate values by calibrating an exponential curve to the two observations.

Using annual data, instead of 10-year averages, substantially increases the number of observations in the analysis. In addition, we can use country fixed effects to control for countries' unobserved heterogeneity, that is, for any country-specific factor as long as it is

constant over time. One of these factors could be the quality of data reporting. As mentioned explicitly by the DFO (2011) “The statistics presented in the DFO are derived from a wide variety of news and governmental sources. The quality and quantity of information available about a particular flood is not always in proportion to its actual magnitude, and the intensity of news coverage varies from nation to nation. In general, news from floods in low-tech countries tends to arrive later and be less detailed than information from 'first world' countries.”

We estimate all the models using Poisson regressions. The dependent variable flood frequency (either as a sum over the 10 years or annually) is a non-negative count variable. Poisson and negative binominal regression are standard estimators in these circumstances. Our preferred estimator is a quasi-maximum likelihood (QML) Poisson model (with fixed effects for the panel regressions) with robust standard errors (Wooldridge 2002, pp. 674-6). Aside from requiring the fixed effect to have a multiplicative effect on the conditional mean, the QML fixed effect Poisson model with robust standard errors does not place any restrictions on conditional distribution of the dependent variable. In addition, it provides consistent estimates of model parameters and their standard errors even if the distribution is characterized by over dispersion (variance is greater than mean) and includes a large number of zeros or exhibits serial correlation (Wooldridge 2002, pp. 674-6).

For the panel models, a random effect Poisson model is also available but Hausman tests reject it in favor of the fixed effect model (chi2 = 42.26 with $P<0.001$). A fixed effects estimate is also preferred conceptually (Wooldridge 2002, pp. 250-1) on the ground that our

sample is not a random draw of countries. Finally, to control for possible time effects, we introduce year dummies in our specification.

5. Results

5.1 Revisiting Bradshaw et al.'s (2007) results

Table 4 shows the results of the estimation of models 1 to 9. The table shows that natural forest cover is negatively associated with the total number of floods between 1990 and 2000 and hence shows existence of the flood-deforestation relationship. The coefficients are statistically significant at a 1 percent level or better and the point coefficients are estimated in a narrow interval (.51-.57). They demonstrate that the results in Bradshaw *et al*. (2007) are not driven by their choice of statistical technique. We also find that non-natural forest cover has the opposite impact. It is associated with an increase in flood frequency. None of the interaction terms is significant.

The choice of preferred model ("best fit model") in Bradshaw *et al*. (2007), Model 6, is also consistent with the patterns of significant of individual coefficients. We take Model 6 as the baseline to assess the robustness of the results in the rest of the analysis.

5.2. Sample effects

In table 5 we show the results of the estimation of the preferred model for different samples. The first column is the baseline. It includes the 41 developing countries that do not drop out because of missing observations. China is excluded, as Bradshaw *et al.* (2007) consider it an outlier. The exclusion of China does not drive the results however, as shown in column 2.

We further extend the sample to include all developing countries. In column 3, natural forest cover is no longer statistically significant to explain total floods. In column 4 we consider all the countries (both developing and developed). Natural forest cover is not significant either.

5.3 Including socioeconomic and institutional factors

We explore the human influence in generating (or reporting) large flood events by including population density, urban population growth, income and corruption into the preferred specification.

Table 6 shows the results. The table shows that the coefficient of natural forest cover is not statistically significant. The only significant effects of the population and urban population growth are for the sample including all the developing countries and all the countries (this is also the larger sample). Income has a positive and significant effect on flood frequency for the sample that includes all countries.

5.4 Panel analysis

Table 7 shows the results of the estimation of the preferred specification using panel data for all the developing countries for the sample period 1990-2000. The first column shows estimates from a pooled Poisson, while in the second, we include country fixed effects. Forest cover is not associated with flood frequency in either specification.

Regarding socioeconomic and institutional factors, income positively affects flood frequency in the pooled Poisson specification. A one percent increase in GDP per capita is associated with a 17 % increase in flood frequency. However, its sign is reversed in the FE Poisson

specification, that captures the within- (rather than the between-) country variation. A one percent increase in GDP per capita lowers flood frequency by 84%.

Corruption is not statistically significant in the pooled Poisson estimates. However, once we account for country-specific unobserved heterogeneity, it becomes statistically significant (at a 10 % level or better) (in column 2), implying that a one unit increase in corruption tends to increase the number of annual floods by 16%. Likewise, urbanization is statistically significant (at less than one percent level) in the pooled Poisson estimates and shows that one percent increase in urban population growth is associated with 15 % increase in the number of annual floods. However, it is no longer significant in the FE model.

We change the sample size and estimation period and re-estimate the model again to see robustness of our results. Column (1) and column (2) of table 8 shows the Poisson estimates after the sample is limited to that in Bradshaw *et al*. (2007) (41 developing countries). Likewise, column (3) and column (4) shows the Poisson estimates after extending the sample to cover all countries. Finally, column (5) and column (6) shows the estimates from the model for all countries after extending the estimation period from 1985-2009.

Natural forest cover exhibits a negative and statistically significant sign in columns (1) and (5). However, once we control for country-specific unobserved factors (column 2, 4, and 6) it ceases being significant and has the opposite sign.

Population tends to have a positive impact on flood frequency, more marked in the pooled models. Income exhibits a positive and significant coefficient in the pooled models of

columns (3) and (5), but it turns negative (and significant for the sample of all countries) after we control for country fixed effects.

An improvement in the corruption indicator (recall that higher values in the scale denote lower corruption) tends to reduce flood frequency. The effect is stronger in the FE models. Finally, urbanization is positively associated with flood frequency but significantly so only in the pooled models.

6. Concluding remarks

We reanalyze the data and models used by Bradshaw *et al.* (2007) and find that using their sample and variables (although a different estimation technique –Poisson regressions), natural forest cover has a robust negative effect on the frequency of large flood events. We show that it is not the econometric technique, nor the exclusion of China from their sample that is driving the results. However, once we extend their sample to cover all developing countries, the coefficient of natural forest cover is no longer statistically significant.

In our analysis, the link between deforestation and floods does not seem robust to the inclusion of variables that account for the "human factor" (population, urban population growth, income and corruption). However given the small number of observations in our analyses we must interpret the results with caution.

Using panel data allows us to increase the sample size substantially and to control for country specific characteristics that do not vary over time (e.g. differences in reporting and other omitted variables as long as these are relatively stable). Forest cover does not have the

expected sign and is not statistically significant. Overall our results indicate that the link between deforestation and the frequency of large floods in developing countries is not robust.

References

Beljaars, A. C. M., P. Viterbo, M. J. Miller and A. K. Betts (1996), "The Anomalous Rainfall Over the United States during July 1993: Sensitivity to Land Surface Parameterization and Soil Moisture Anomalies" *Monthly Weather Review*, 124, (3): 362-383.

Bradshaw, J. A., N. S. Sodhi, K. S. Peh and B. W. Brook (2007), "Global Evidence that Deforestation Amplifies Flood Risk and Severity in the Developing World" *Global Change Biology*, 2007, (13): 2379-2395.

Bruijnzeel, L. (2004), "Hydrological Functions of Tropical Forests: Not Seeing the Soil for the Trees?" *Agriculture, Ecosystems & Environment*, 104, (1): 185-228.

Calder, I. R. (2007), "Forests and Water--Ensuring Forest Benefits Outweigh Water Costs" *Forest Ecology and Management*, 251, (1-2): 110-120.

Cassardo, C., G. Balsamo, C. Cacciamani, D. Cesari, T. Paccagnella and R. Pelosini (2002), "Impact of Soil Surface Moisture Initialization on Rainfall in a Limited Area Model: A Case Study of the 1995 South Ticino Flash Flood" *Hydrological Processes*, 16, (6): 1301-1317.

CIFOR (Center for International Forestry Research)/FAO (Food and Agriculture Organization of the United Nations) (2005) "Forests and Foods Drowning in Fiction or Thriving on Facts?", *RAP Publ i cat ion 2005/03*.

Congleton, R. D. (2006), "The Story of Katrina: New Orleans and the Political Economy of Catastrophe" *Public Choice*, 127, (1): 5-30.

CRED/OFDA (Center for Research of the Epidemiology of Disasters/Office of Foreign Disaster Assistance) (2011), "International Disaster Database", www.emdat.be (Accessed June 12, 2011).

Deacon, R. (2003), "Dictatorship, Democracy, and the Provision of Public Goods" University of California at Santa Barbara, Economics Working Paper Series, Department of Economics, UC Santa Barbara.

DFO (Dartmouth Flood Observatory) (2011), "Space-based Measurement of Surface Water", http://floodobservatory.colorado.edu/index.html (Accessed Jan. 16, 2011).

FAO (Food and Agriculture Organization of the United Nations) (2010), "Global Forest Resources Assessment 2010", http://www.fao.org/forestry/fra/fra2010/en/ (Accessed April 12, 2011).

Ferreira, S. (2010), "An Analysis of the Determinants of Flood Damages" 2011 Annual Meeting, February 5-8, Southern Agricultural Economics Association.

Kahn, M. E. (2005), "The Death Toll from Natural Disasters: the Role of Income, Geography, and Institutions" *Review of Economics and Statistics*, 87, (2): 271-284.

Kisslinger, C. (1976), "A Review of Theories of Mechanisms of Induced Seismicity" *Engineering Geology*, 10, (2-4): 85-98.

Konrad, C. P. (2005) "Effects of Urban Development on Floods", *U.S. Geological Survey Fact Sheet 076-03*.

Myers, N. (1986), "Environmental Repercussions of Deforestation in the Himalayas" *Journal World Forest Resource Management* 2: 63-75.

Nunn, N. and D. Puga (2010), "Ruggedness: The Blessing of Bad Geography in Africa" *Review of Economics and Statistics*, Forth coming.

PRS (Political Risk Services) (2011), "International Country Risk Guide", https://www.prsgroup.com/prsgroup_shoppingcart/cdwizard5.aspx (Accessed Feb. 5, 2011).

Reed, D. W. (2002), "Reinforcing Flood-Risk Estimation" *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 360, (1796): 1373-1387.

TCCCR (Tyndall Centre for Climate Change Research) (2011), "Climate Data", http://www.cru.uea.ac.uk/cru/data/hrg/timm/cty/obs/TYN_CY_1_1.html (Accessed Feb. 10, 2011).

Tollan, A. (2002), "Land-use Change and Floods: What Do We Need Most, Research or Management?" *Water Science and Technology*: 183-190.

USDA (US Department of Agriculture) (2011), "Soil Moisture Regimes Map", http://soils.usda.gov/use/worldsoils/mapindex/smr.html (Accessed Feb. 5, 2011).

Van Dijk, A. I. J. M. and R. J. Keenan (2007), "Planted Forests and Water in Perspective" *Forest Ecology and Management*, 251, (1-2): 1-9.

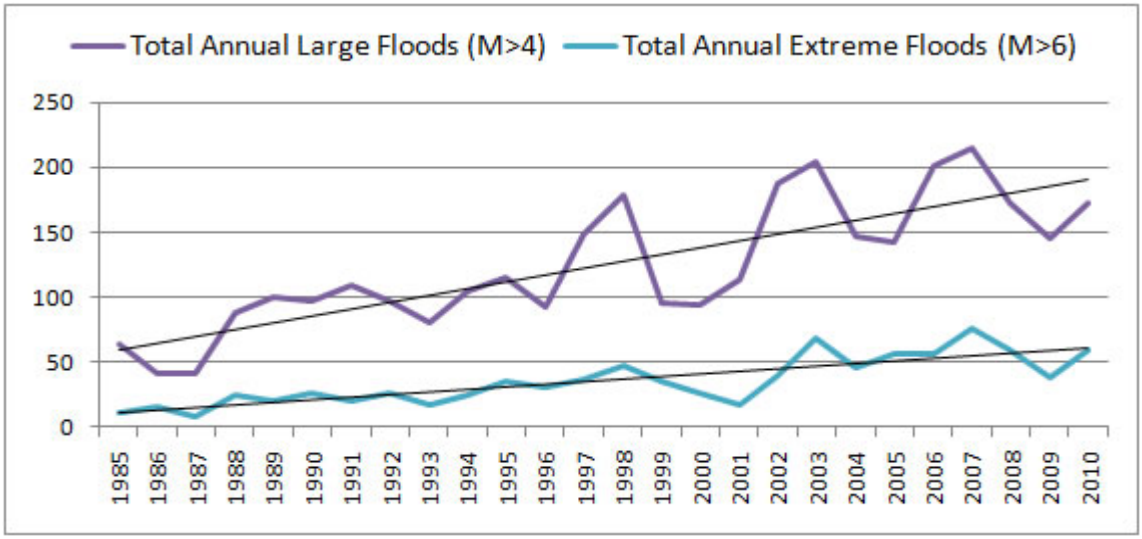
Van Dijk, A. I. J. M., M. Van Noordwijk, I. R. CaldeR, S. L. A. Bruijnzeel, J. Schellekens and N. A. Chappell (2009), "Forest-Flood Relation Still Tenuous - Comment on 'Global Evidence that Deforestation Amplifies Flood Risk and Severity in the Developing World' by C. J. A. Bradshaw, NS Sodi, K. S.-H. Peh and BW Brook" *Global Change Biology*, 15, (1): 110-115.

WDI (World Development Indicators) (2010), "World Development Indicators", http://data.worldbank.org/data-catalog/world-development-indicators/wdi-2010 (Accessed March 12, 2011).

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

WRI (World Resources Institute) (2011), "Earth Trends: Environmental Information", http://earthtrends.wri.org/ (Accessed Feb. 12, 2011).

Figure 1: Total annual large floods (1985-2010)



Notes: M is flood magnitude, computed as Log (duration in days * severity * affected area in square km). Severity can take the values 1, 1.5 and 2. (Floods are divided into three severity classes depending on their estimated recurrence interval. Class 1 floods have a 10-20 year-long reported interval between similar events, class 1.5 has a 20-100 year recurrence interval, and class 2 has a recurrence interval greater than 100 years.) Source: DFO (2011)

Table 1: Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
Annual number of floods	0.51	1.40	0	25
Country area, sq km	814475.40	1999742	50	1.64E+07
Annual rainfall, mm	1128.38	788.69	35.4	3355.2
Soil moisture: arid=1	0.29	0.45	0	1
Soil moisture: sub-humid=1	0.25	0.43	0	1
Average slope (%)	3.95	3.34	0.03	17.60
Area of degraded land (%)	173601	435853.2	0	2700413
Total population	3.60E+07	1.25E+08	57544.84	1.26E+09
Annual natural forest cover, sq km	209533.50	659923.50	0	7939090
Annual non-natural forest cover, sq km	14781.51	49303.45	0	543940
GDP PC PPP @ constant 2005 $	8260.55	9682.88	150.80	48397.6
Corruption	3.27	1.30	0	6
Urban population growth (%)	2.42	2.21	-5.731994	20.01

Sources: Flood data from DFO (2011); WDI (2010) for GDP per capita PPP, country area, total population and urban population growth data; FAO (2010) for natural and non-natural forest data; USDA (2011) for soil moisture data; Nunn & Puga (2010) for slope data; PRS (2011) for data related to corruption; TCCCR (2011) for rainfall data; and WRI (2011) for data of degraded land.

Table 2: Correlation matrix

	Total floods	Area	Annual rainfall	Soil moisture: arid=1	Soil moisture: sub-humid=1	Average slope (%)	Degraded land (%)	Total population	Natural forest cover, sq km	Non natural forest cover, sq km	GDP PC PPP	Corruption	Urban population growth
Total floods	1												
Area	0.60	1											
Annual rainfall	-0.04	-0.18	1										
Soil moisture: arid=1	-0.03	0.11	-0.63	1									
Soil moisture: sub-humid=1	-0.08	-0.12	0.08	-0.38	1								
Average slope (%)	0.09	-0.09	-0.01	0.10	-0.28	1							
Degraded land (%)	0.62	0.85	-0.08	-0.08	-0.01	-0.02	1						
Total population	0.54	0.59	-0.06	-0.10	0.06	0.06	0.80	1					
Natural forest cover, sq km	0.38	0.80	0.05	-0.09	-0.06	-0.14	0.71	0.29	1				
Non natural forest cover, sq km	0.65	0.67	-0.11	-0.12	-0.11	0.07	0.75	0.81	0.36	1			
GDP PC PPP	0.06	0.09	-0.12	0.05	-0.36	0.23	0.05	-0.07	0.07	0.10	1		
Corruption	-0.01	0.04	-0.13	-0.07	-0.19	0.26	0.04	-0.05	0.04	0.09	0.52	1	
Urban population growth	0.05	0.03	0.04	0.13	0.19	-0.18	-0.02	0.05	0.00	-0.05	-0.36	-0.40	1

Table 3: Variables used by Bradshaw *et al*. (2007) in different model

Model no.	Variables used
1	AR, RN, SL, DG, NFC, NFL, NNFC, NFC*NFL, SL*NFC, SL*NFL
2	AR, RN, SL, DG, NFC, NFL, NNFC
3	AR, RN, SL, DG, NFC, NFL, NNFC, NFC*NFL
4	AR, RN, SL, DG, NFC, NFL, NNFC, SL*NFC
5	AR, RN, SL, DG, NFC, NFL, NNFC, SL*NFL
6	AR, RN, SL, DG, NFC, NNFC
7 & 8	AR, RN, SL, DG, NFC, NFL
9	AR, RN, SL, DG

Abbreviation used: AR – country area; RN – average annual rainfall; SL- average slope; DG – total degraded land; NFC – natural forest cover; NNFC – non-natural forest cover; NFL – natural forest loss and * denotes interaction between variables.

Table 4: Analysis of Bradshaw *et al*. (2007) results

	Model definitions as in Table 3							
VARIABLES	Model 9	Models 7&8	Model 6	Model 5	Model 4	Model 3	Model 2	Model 1
Ln(country area, sq km)	0.0226	0.789***	0.775***	0.716**	0.705**	0.747**	0.733**	0.744**
	(0.153)	(0.298)	(0.301)	(0.325)	(0.314)	(0.294)	(0.305)	(0.339)
Ln(average annual rainfall over 1900-2000, mm)	0.537	0.398	0.531*	0.279	0.241	0.255	0.263	0.282
	(0.431)	(0.243)	(0.308)	(0.279)	(0.275)	(0.281)	(0.273)	(0.281)
Soil moisture: arid=1	0.477	-0.473	-0.434	-0.516	-0.505	-0.504	-0.562	-0.471
	(0.576)	(0.559)	(0.474)	(0.520)	(0.512)	(0.464)	(0.482)	(0.520)
Soil moisture: sub-humid=1	0.131	-0.105	-0.107	-0.104	-0.0970	-0.0182	-0.128	0.0135
	(0.263)	(0.222)	(0.238)	(0.198)	(0.218)	(0.215)	(0.206)	(0.209)
Average slope (%)	0.0736	0.0236	-0.00105	-0.0174	-0.0268	0.00332	-0.0250	0.0182
	(0.0535)	(0.0432)	(0.0436)	(0.0440)	(0.0462)	(0.0524)	(0.0453)	(0.0632)
Ln(degraded land, sq km)	0.465***	0.222*	0.0255	0.0690	0.0655	0.0457	0.0621	0.0493
	(0.142)	(0.120)	(0.137)	(0.136)	(0.133)	(0.130)	(0.130)	(0.144)
Ln(average annual natural forest cover over 1990-2000, sq km)		-0.572***	-0.511***	-0.515**	-0.550***	-0.569***	-0.532***	-0.549***
		(0.212)	(0.195)	(0.212)	(0.203)	(0.200)	(0.198)	(0.206)
Average annual change in natural forest cover over 1990-2000 (%)		0.145**		0.0710	0.118*	0.140*	0.118*	0.0924
		(0.0581)		(0.105)	(0.0711)	(0.0725)	(0.0690)	(0.122)
Ln(average annual non-natural forest cover over 1990-2000, sq km)			0.226***	0.161**	0.163**	0.157**	0.168**	0.149*
			(0.0730)	(0.0752)	(0.0754)	(0.0769)	(0.0733)	(0.0800)
Interaction: slope & average change in natural forest cover over 1990-2000				0.0213				0.0238
				(0.0328)				(0.0352)
Interaction: slope & average natural forest cover over 1990-2000					8.38E-08			-4.40E-08
					(1.29E-07)			(2.30E-07)
Interaction: average natural forest cover over 1990-2000 & its annual change						-1.83E-07		-2.19E-07
						(1.49E-07)		(2.38E-07)
Constant	-8.087**	-7.033***	-8.004***	-5.295**	-4.475	-4.711*	-5.142**	-5.132*
	(3.678)	(2.160)	(2.957)	(2.493)	(2.752)	(2.737)	(2.510)	(2.905)
Observations	51	42	41	39	39	39	39	39

Poisson estimates. Dependent variable is total number of floods between 1990 and 2000. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

Table 5: Robustness of Bradshaw *et al.*'s (2007) results to sample definition

VARIABLES	Baseline	+ China	All Developing	All
Ln(country area, sq km)	0.775***	0.881***	0.564**	0.598***
	(0.301)	(0.261)	(0.232)	(0.156)
Ln(average annual rainfall over 1900-2000, mm)	0.531*	0.378	0.235	0.211
	(0.308)	(0.308)	(0.393)	(0.302)
Soil moisture: arid=1	-0.434	-0.833***	-0.125	-0.252
	(0.474)	(0.276)	(0.291)	(0.252)
Soil moisture: sub-humid=1	-0.107	-0.291	0.0329	-0.0913
	(0.238)	(0.227)	(0.271)	(0.272)
Average slope (%)	-0.00105	0.00167	0.0734**	0.0737***
	(0.0436)	(0.0421)	(0.0299)	(0.0277)
Ln(degraded land, sq km)	0.0255	0.00580	0.00248	0.130
	(0.137)	(0.135)	(0.105)	(0.0858)
Ln(average annual natural forest cover over 1990-2000, sq km)	-0.511***	-0.610***	-0.0907	-0.142
	(0.195)	(0.159)	(0.185)	(0.123)
Ln(average annual non-natural forest cover over 1990-2000, sq km)	0.226***	0.251***	0.170**	0.109*
	(0.0730)	(0.0719)	(0.0787)	(0.0622)
Constant	-8.004***	-6.978**	-7.738**	-8.358***
	(2.957)	(3.198)	(3.320)	(2.618)
Observations	41	42	66	82

Poisson estimates. Dependent variable is total number of floods between 1990 and 2000. Baseline is Model 6 in Table 4.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Controlling socioeconomic and institutional variables in Bradshaw *et al*. (2007) and robustness of the results to sample definition

VARIABLES	Bradshaw *et al* sample	+ China	All developing	All
Ln(country area, sq km)	0.291	0.518	0.0906	0.413***
	(0.466)	(0.349)	(0.194)	(0.129)
Ln(average annual rainfall over 1900-2000, mm)	0.508**	0.303	0.0114	0.110
	(0.253)	(0.244)	(0.181)	(0.136)
Soil moisture: arid=1	0.297	-0.353	-0.127	-0.398**
	(0.574)	(0.434)	(0.207)	(0.163)
Soil moisture: sub-humid=1	-0.0543	-0.321	-0.260*	-0.213
	(0.199)	(0.197)	(0.157)	(0.143)
Average slope (%)	0.0721	0.0751	0.107***	0.106***
	(0.0571)	(0.0553)	(0.0312)	(0.0256)
Ln(degraded land, sq km)	0.0539	0.0195	-0.107	0.0675
	(0.124)	(0.110)	(0.0905)	(0.0759)
Ln(average annual natural forest cover over 1990-2000, sq km)	-0.112	-0.303	0.122	-0.115
	(0.325)	(0.245)	(0.115)	(0.0738)
Ln(average annual non-natural forest cover over 1990-2000, sq km)	0.0522	0.102	0.0100	-0.0523
	(0.0904)	(0.105)	(0.0563)	(0.0508)
Ln(average annual total population over 1990-2000)	0.253	0.226	0.479***	0.355***
	(0.238)	(0.215)	(0.114)	(0.0848)
Ln(average GDP PC PPP over 1990-2000)	0.0223	0.0328	0.105	0.304***
	(0.125)	(0.116)	(0.0903)	(0.0725)
Average corruption over 1990-2000	-0.119	-0.108	0.0650	0.0217
	(0.159)	(0.143)	(0.0988)	(0.0793)
Average annual urban population growth over 1990-2000	0.109	0.120	0.162***	0.244***
	(0.114)	(0.119)	(0.0530)	(0.0490)
Constant	-9.983***	-8.711**	-9.533***	-12.80***
	(3.039)	(3.737)	(1.789)	(1.663)
Observations	35	36	52	68

Note: Poisson estimates. Dependent variable is total number of floods a country experienced over 1990-2000. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Determinants of flood frequency in the developing countries (1990-2000)

VARIABLES	Pooled Poisson	FE Poisson
Ln(country area, sq km)	0.200	
	(0.193)	
Ln(annual rainfall, mm)	0.180	2.373***
	(0.215)	(0.442)
Soil moisture: arid=1	-0.0701	
	(0.249)	
Soil moisture: sub-humid=1	-0.258	
	(0.169)	
Average slope (%)	0.137***	
	(0.0367)	
Ln(degraded land, sq km)	-0.103	
	(0.0962)	
Ln(total population)	0.507***	0.882
	(0.117)	(2.199)
Ln(annual natural forest cover, sq km)	0.0284	2.970
	(0.121)	(2.439)
Ln(annual non-natural forest cover, sq km)	0.00612	-0.295
	(0.0637)	(0.560)
Ln(GDP PC PPP at 2005 $)	0.177**	-0.843*
	(0.0848)	(0.453)
Corruption	-0.0616	-0.159*
	(0.0834)	(0.0958)
Urban population growth	0.154***	0.103
	(0.0533)	(0.113)
Constant	-14.43***	
	(2.278)	
Year effect	Yes	Yes
Observations	587	480
Number of id		48

Note: Poisson estimates. Dependent variable is total yearly floods. Robust Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Robustness of the results

VARIABLES	Bradshaw *et al*. sample		All countries		All-period (1985-2009)	
	Pooled Poisson	FE Poisson	Pooled Poisson	FE Poisson	Pooled Poisson	FE Poisson
Ln(country area, sq km)	0.423***		0.332***		0.471***	
	(0.153)		(0.109)		(0.0688)	
Ln(annual rainfall, mm)	0.677***	1.227***	0.163	2.439***	0.451***	1.887***
	(0.132)	(0.335)	(0.160)	(0.351)	(0.0944)	(0.279)
Soil moisture: arid=1	0.151		-0.154		0.0248	
	(0.165)		(0.190)		(0.134)	
Soil moisture: sub-humid=1	-0.331***		-0.303*		-0.362***	
	(0.119)		(0.171)		(0.110)	
Average slope (%)	0.113***		0.0714***		0.0549***	
	(0.0252)		(0.0252)		(0.0162)	
Ln(degraded land, sq km)	0.0111		0.0842		0.00890	
	(0.0790)		(0.0758)		(0.0508)	
Ln(total population)	0.315***	1.933**	0.492***	1.838	0.470***	1.291*
	(0.0858)	(0.959)	(0.0880)	(1.679)	(0.0585)	(0.773)
Ln(annual natural forest cover, sq km)	-0.161*	1.223	-0.0943	3.336	-0.109***	0.676
	(0.0939)	(0.814)	(0.0764)	(2.154)	(0.0408)	(0.674)
Ln(annual non-natural forest cover, sq km)	0.0505	0.121	-0.0704	-0.140	-0.0782**	0.238
	(0.0435)	(0.322)	(0.0528)	(0.529)	(0.0356)	(0.324)
Ln(GDP PC PPP at 2005 $)	0.00265	-0.276	0.366***	-0.850**	0.335***	-0.346
	(0.0693)	(0.222)	(0.0787)	(0.381)	(0.0637)	(0.223)
Corruption	-0.162***	-0.171*	-0.0814	-0.156**	-0.0354	-0.138**
	(0.0427)	(0.0902)	(0.0641)	(0.0748)	(0.0432)	(0.0683)
Urban population growth	0.136***	0.149	0.232***	0.0637	0.218***	0.110
	(0.0433)	(0.0924)	(0.0515)	(0.0966)	(0.0416)	(0.0783)
Constant	-14.55***		-17.33***		-18.85***	
	(1.517)		(2.078)		(1.341)	
Year effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	718	720	795	623	1,607	1,413
Number of id		41		63		80

Note: Poisson estimates. Dependent variable is total yearly floods. Robust Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix: Table 1

List of countries used by Bradshaw *et al*. (2007) (56 developing countries)
Angola, Cameroon, Botswana, China, Central African Republic, Chad, Congo, Democratic Republic of Congo, Ethiopia, Ghana, Kenya, Malawi, Mozambique, Nigeria, Republic of South Africa, Senegal, Sudan, Tanzania, Togo, Uganda, Zambia, Zimbabwe, Bangladesh, Cambodia, India, Indonesia, Laos, Malaysia, Myanmar, Philippines, Papua New Guinea, Sri Lanka, Thailand, Vietnam, Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, EL Salvador, Guatemala, Guyana, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Trinidad and Tobago, Uruguay, Venezuela.
List of the developing countries used in our analysis (66 developing countries)
Albania, Algeria, Argentina, Bangladesh, Belarus, Benin, Bolivia, Bosnia, Brazil, Burkina Faso, Cambodia, Central African Republic, Chad, Chile, China, Columbia, Congo Republic, Costa Rica, Cuba, EL Salvador, Ethiopia, Georgia, Ghana, Guatemala, Haiti, India, Indonesia, Iran, Jamaica, Kazakhstan, Kenya, Laos, Liberia, Malawi, Malaysia, Mauritania, Mexico, Mongolia, Morocco, Mozambique, Myanmar, Nepal, Niger, Nigeria, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Poland, Romania, Senegal, Somalia, South Africa, Sri Lanka, Sudan, Syria, Tajikistan, Thailand, Togo, Tunisia, Turkey, Uganda, Ukraine, Vietnam, Zimbabwe. However, some countries dropped out in our FE estimates because of lack of sufficient observation within group.
List of all countries used in our analysis (82 countries)
Albania, Algeria, Argentina, Australia, Bangladesh, Belarus, Benin, Bolivia, Bosnia, Brazil, Burkina Faso, Cambodia, Canada, Central African Republic, Chad, Chile, China, Czech Republic Columbia, Congo Republic, Costa Rica, Cuba, EL Salvador, Ethiopia, Fiji, France, Germany, Georgia, Ghana, Greece, Guatemala, Haiti, India, Indonesia, Iran, Israel, Italy, Jamaica, Japan, Kazakhstan, Kenya, Laos, Liberia, Malawi, Malaysia, Mauritania, Mexico, Mongolia, Morocco, Mozambique, Myanmar, Nepal, New Zealand, Niger, Nigeria, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Poland, Poland, Portugal, Romania, Senegal, Slovakia, Somalia, Spain, South Africa, Sri Lanka, Sudan, Syria, Tajikistan, Thailand, Togo, Trinidad & Tobago, Tunisia, Turkey, Uganda, United Kingdom, United States of America, Ukraine, Vietnam, Zimbabwe. However, some countries dropped out in our FE estimates because of lack of sufficient observation within group.