Natural Hazards and Economic Losses: Why Correcting Sample Selection Matters*

Dahye Lina Song† and Christine Choirat‡

May 4, 2015

Abstract

Economic losses from natural disasters vary by countries, and it has been hypothesized that institutional, political, and other national conditions and policies all play a role in determining the severity of loss. Many empirical studies for understanding the determinants of disaster losses, however, suffer from endogeneity and selection bias, which can potentially make their results method-dependent. To demonstrate, we investigate the relationship between disaster propensity, wealth, and economic loss from a panel data collected by [Neumayer et al., 2014].

*Our thanks to the students of GOV2001 for reading this draft, and Professor King, Solé, and Stephen for helpful advices and feedbacks.

†PhD Program in Health Policy, 14 Story Street, Harvard University, Cambridge MA 02138; dsong@fas.harvard.edu

‡Institute for Quantitative Social Science, 1737 Cambridge Street, Harvard University, Cambridge MA 02138; cchoirat@iq.harvard.edu

We first demonstrate that the original data is subject to endogeneity and selection bias, reconstruct the dataset, and apply Heckman correction. The bias-corrected estimated impact of disaster propensity changes direction from the original result by [Neumayer et al., 2014] — countries that experience more frequent disasters tend to suffer from greater economic damage, holding everything else equal. We suggest that disaster propensity could be an indicator of vulnerability, or a sign of insufficient prevention and mitigation measures. Although we cannot provide any definitive explanation for the phenomenon, our result shows that correcting selection bias matters when dealing with natural disasters data. For future work, a more sophisticated construction of the latent propensity variable and the application of quantile regression for endogenous selection models could broaden our understanding.

1 Introduction

In 2013, 337 disasters related to natural hazards were reported worldwide, causing $22,452$ deaths and $119,217$ million USD of disaster estimated damage [IFRC, 2014]. Global society has long been influenced by natural hazards, and there is a growing consensus among the scientific community that climate change is exacerbating hot and precipitation extremes [Fischer and Knutti, 2015]. Furthermore, the number of people exposed to strong earthquakes in both developed and developing countries has grown exponentially over the past 60 years, contributing to the increase of quake vulnerable populations [Tucker, 2013].

Some countries or communities have higher economic damage than others that face similar magnitude of disaster. Many hypothesize that institutional, political, and other national conditions play a role in determining disaster impacts [Kousky, 2014]. Earlier studies suggest that richer nations suffer less damage because economic development and high quality institutions provide implicit insurance against nature's shocks [Kahn, 2005, Toya and Skidmore, 2007, Kellenberg and Mobarak, 2008]. More recently, some studies have started investigating the role of disaster hazard. [Schumacher and Strobl, 2011] shows that, for two countries with equal wealth, the one with lower hazard rates should invest less in mitigation and then could conceivably suffer more damages when an event occurs. [Neumayer et al., 2014] also investigates the role of disaster propensity in mitigation of economic loss. Using one of the most extensive data on losses from natural disasters from Munich Re, a global reinsurer, it argues that countries more prone to disasters will invest more in mitigation and prevention efforts and thus are more likely to have less damages.

Yet, it is still difficult to understand through what mechanism hazard affects disaster loss. It is not clear whether the hazard is just a proxy for other key variables that contribute to mitigating losses, or whether the hazard actually provides individuals and governments an incentive to take active mitigation and prevention measures. Many of the studies suffer from bias and endogeneity problems, because of the nature of available disaster datasets. These datasets are often subject to difficulties in measurement, inconsistent

inclusion criteria, and various types of reporting biases [Wirtz et al., 2014, Kron et al., 2012].

In this paper, we demonstrate that correcting selection bias is crucial in establishing a meaningful empirical evidence between disaster propensity and economic loss. Adapting the data and study of [Neumayer et al., 2014], we argue that detecting a sign of government's role through a disaster loss dataset is difficult, and sometimes misleading. We first show that endogeneity and selection bias exist in [Neumayer et al., 2014]'s analysis due to the definition of disaster propensity variable and the limitations from the construction of dataset. Then we apply Heckman correction [Heckman, 1979, Toomet and Henningsen, 2008] to show that the bias corrected effect of disaster propensity changes the direction from the original result. We conclude with a cautionary message that the evidence is not clear on whether governments invest more in disaster prevention and damage mitigation when the expected damage of a potential disaster increases.

2 Disaster propensity and selection bias

2.1 Dataset construction

The original dataset from [Neumayer et al., 2014] is an unbalanced panel data of 201 countries for the period of 1980 to 2008. It consists of three types of natural disasters – earthquakes, tropical cyclones, and floods – which account for about 70% of total worldwide natural disasters in Mu-

nich Re dataset. To measure the propensity of quake and tropical cyclone events, [Neumayer et al., 2014] uses the Richter scale and top wind speed information from the Munich Re database. It uses the precipitation measures from [Willmott and Matsuura, 2011] for flood propensity by matching the geographical location of the disaster center. According to [Neumayer et al., 2014], these measures capture the main destructive forces of the respective hazard events and thus is the best available proxy in the absence of data for other relevant magnitude variables.

The disaster hazard magnitude variable was constructed by aggregating the data from individual disaster event to the country-year level. Disaster propensity for a country i is defined in [Neumayer et al., 2014, p. 13] as the sum of all hazard magnitudes such that

$$\mathsf{propensity}_i = \sum_{t=1980}^{2008} \mathsf{hazard\ magnitude}_{i,t}. \tag{1}$$

These variables capture the two key dimensions of propensity, frequency and magnitude, since they are systematically higher when a country experiences more frequent hazards of a certain type and when it experiences stronger hazard events. To estimate the effect of disaster propensity on economic loss, [Neumayer et al., 2014] quantile-regresses log economic loss on four explanatory variables: log disaster hazard magnitude, log disaster propensity, log per capita income of country (GDP per capita), and log gross domestic product (GDP) of country.

2.2 Dataset inspection

Figure 1 shows the trend of log losses over time for four representative countries of varying GDP levels. Every dot corresponds to a country-year disaster loss and lines correspond to the trend of log losses over time per disaster type. The trend for the whole range of countries is displayed in Figure 4 and Figure 5 in the Appendix.

Between the two countries with high GDP, Japan and the United States, the quake propensity of Japan is much higher than that of the U.S. (29.52 vs 28.58). From the plot, however, we see a steeper increasing trend of log losses for Japan than that of the U.S., suggesting potential insufficient mitigation effort in response to higher hazard. From inspecting the two countries with lower GDP, China and India, we see that although the quake propensity of India (31.21) is higher than the one of China (28.21), more disasters are reported in China than India (28 vs 20), suggesting some potential missing observations from India. Rising trends in observations of all disaster types in China also suggest that some of it might be attributable to higher reporting rates over time.

2.3 Dataset and endogeneity

We believe the model proposed by [Neumayer et al., 2014] provides a biased estimate of the relationship between disaster propensity and disaster losses, and one of the reasons is that there is an endogeneity issue from the way data was constructed. To estimate the unbiased effect of propensity on economic

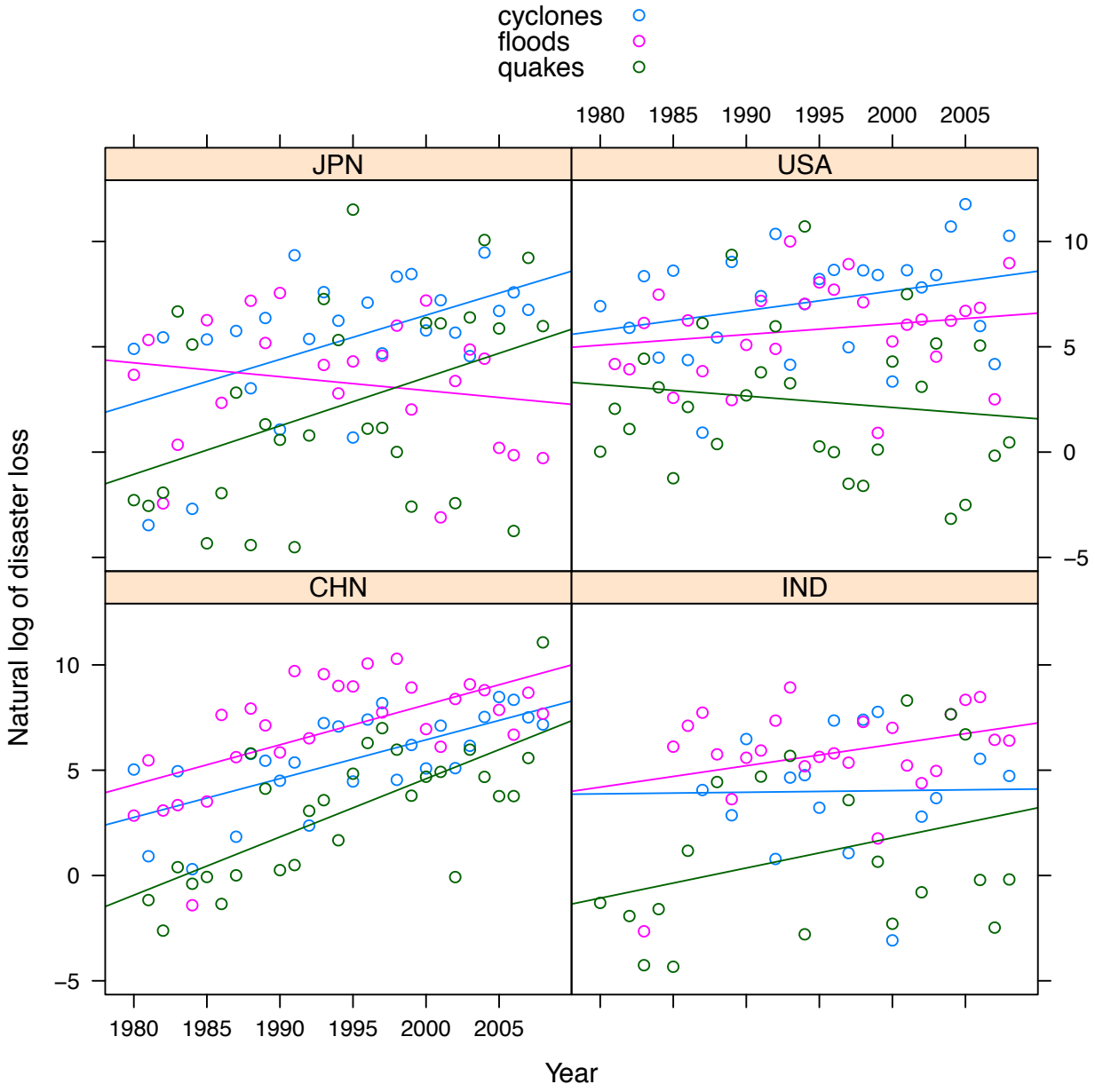


Figure 1: Log losses by disaster type for period 1980-2008 for Japan, USA, China and India

loss, the propensity variable should be exogenous. However, as we can see from Equation 1, propensity is correlated with disaster magnitude by construction. Such problem is conspicuous especially for the countries with only one observation between $1980 - 2009$, since the propensity and hazard magnitude are identical. There are 16 countries with only one observation and 14 countries with two observations, which contributes to the high correlation between disaster propensity and disaster magnitude.

2.4 The evidence of selection bias

In addition to endogeneity, the dataset from [Neumayer et al., 2014] suffers from selection bias. Because country-years with no damages are excluded from the sample, the observation only includes the country-year in which at least one disaster occurred. We believe that the observation is not randomly assigned, however, since the probability of a disaster happening is correlated with the disaster propensity variable constructed by [Neumayer et al., 2014], if that measure is actually serving its definition. More formally, if there is a latent variable D^* for each country-year such that the disaster is observed only when $D^* > 0$, then the dataset does not include the country-year in which $D^* < 0$. Thus, the naive estimate of the effect of propensity on economic loss is only among the countries that are more likely to encounter disasters and/or have them reported to Munich Re.

In addition, the data collection process of Munich Re might also exacerbate selection bias. [Wirtz et al., 2014, Kron et al., 2012] point out that Mu-

nich Re's dataset must be handled with caution, as there are many challenges associated with data acquisition and data management for disaster events and damages. Reporting bias is especially plausible, as [Neumayer et al., 2014] points out, since smaller disasters from the early years of data collection might have escaped Munich Re's attention, and disaster events in the developing world are likely to be underreported. These grant us more reasons to believe that the selection of sample was not random. If some observations are selectively missing for country-specific characteristics, the probability of observation would be correlated with GDP and GDP per capita. If some observations are selectively missing because of small disaster scale, the probability of observation would be correlated with disaster propensity, since it is a proxy for the intensity of disaster as well as frequency.

3 Revisiting [Neumayer et al., 2014]

3.1 Creating a balanced panel

To address potential selection bias, we create a balanced panel of country-year by first adding the country-year with no observation into the dataset. We also create a variable called $\mathsf{observation}$, where $\mathsf{observation}_{i,t} = 1$ if disaster occurred in country i in year t, and $= 0$ otherwise. For all country-years such that $\mathsf{observation}_{i,t} = 0$, $\mathsf{propensity}_{i,t} = \mathsf{propensity}_i$, since the unobserved country-years have zero disaster hazard magnitude and thus do not contribute to the disaster propensity. We collect the GDP and GDP per capita

information from World Bank national accounts data, and OECD National Accounts data files [WB, 2014]. To ensure the match of GDP and GDP per capita between the original and the new data, we compare the values for country-year entries with $\mathsf{observation}_{i,t} = 1$. When there are some discrepancies, we take the average of the two sources.

3.2 Correcting selection bias

We use Heckman correction by [Heckman, 1979] to correct selection bias. We use the two-stage estimation method, where in the first stage we model the probability of a country-year having an observation of disaster, $\mathsf{observation}_{i,t} = 1$, as a probit model with the following covariates: log disaster hazard magnitude, log per capita income of country, and log gross domestic product of country. We include these covariates because of their potential correlation with observation, as we discussed in Section 2.4. Specifically, we assume that the model looks like

$$\Pr(\mathsf{observation}_{i,t} = 1 | X_{i,t}) = \Phi(X'_{i,t}\beta) \tag{2}$$

where $X_{i,t}$ are the covariates and Φ is the cumulative distribution function of the standard normal distribution.

In the second stage, we incorporate the predicted probability of observation as an additional explanatory variable. The conditional expectation of log economic loss given the country-year has an observation is

$$\mathrm{E}[\ln\ \mathsf{loss}_{i,t}|Z_{i,t}, \mathsf{observation}_{i,t} = 1] = Z_{i,t}\beta + \rho\sigma_u\lambda(X_{i,t}\gamma) \tag{3}$$

where $Z_{i,t}$ is the original covariate vector, and ρ, σ_u, λ and γ can be estimated under the assumption that the error terms are jointly normal [Sartori, 2003, Section 2.2]. We are ultimately interested in $\mathrm{E}[\ln\ \mathsf{loss}_{i,t}|Z_{i,t}]$, which can be estimated by combining conditional expectations of $\ln\ \mathsf{loss}$ given $\mathsf{observation}_{i,t} = 1$ and $\mathsf{observation}_{i,t} = 0$.

4 Results and discussion

4.1 Probability of observing a disaster

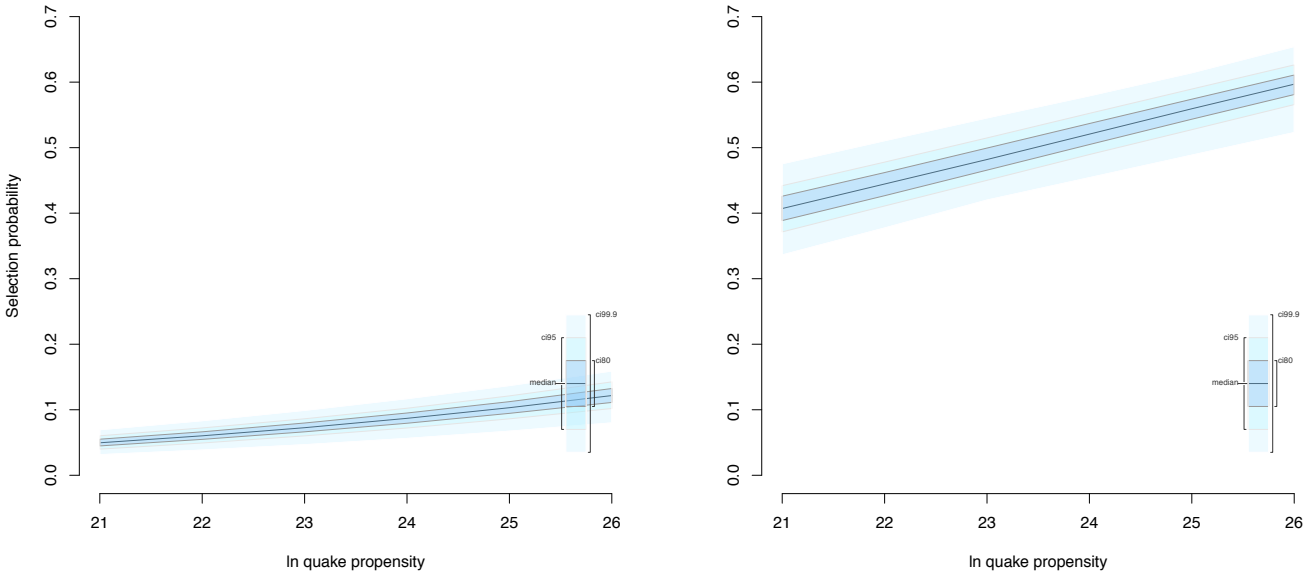
Using R's heckit function [Toomet and Henningsen, 2008], we estimate the sample-selection corrected coefficients. Table 1 presents estimates of the effect of disaster propensity, GDP, and GDP per capita on the probability of observing a disaster in a country-year. As expected, disaster propensity for earthquakes, cyclones, and floods all predict observation (magnitude 0.1, 0.46, and 0.19, respectively), indicating that countries with higher disaster propensity are more likely to encounter disasters. Interestingly, GDP per capita for quakes and floods are negatively associated with the likelihood of observation, whereas GDP increases with the likelihood of observation for all three disaster types. One plausible explanation might be that the level of individual's wealth effectively plays a role in prevention. For example,

Table 1: Economic loss from natural disasters - Selection equation

	Dependent variable: selection probability		
	Disaster type:		
	quakes	cyclones	floods
ln quake propensity	0.10^{***}		
	(0.01)		
ln tropical cyclone propensity		0.46^{***}	
		(0.03)	
ln flood propensity			0.19^{***}
			(0.02)
ln per capita income of country	-0.09^{***}	-0.01	-0.27^{***}
	(0.02)	(0.03)	(0.02)
ln gross domestic product of country	0.23^{***}	0.10^{***}	0.35^{***}
	(0.01)	(0.02)	(0.02)
Constant	-7.80^{***}	-10.00^{***}	-7.60^{***}
	(0.34)	(0.57)	(0.28)
Observations	2,964	1,596	4,059
R^2	0.34	0.37	0.34
Adjusted R^2	0.33	0.37	0.34
ρ	0.94	0.88	1.00
Inverse Mills Ratio	4.10^{*} (2.20)	3.50^{*} (2.10)	4.50^{***} (0.81)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2: Economic loss from natural disasters - Outcome equation

	Dependent variable: natural log of disaster loss		
	Disaster type:		
	quakes	cyclones	floods
ln quake hazard magnitude	0.49***		
	(0.03)		
ln quake propensity	0.12		
	(0.15)		
ln tropical cyclone hazard magnitude		1.40***	
		(0.13)	
ln tropical cyclone propensity		1.10	
		(0.69)	
ln flood hazard magnitude			0.58***
			(0.04)
ln flood propensity			0.45***
			(0.13)
ln per capita income of country	−0.40**	−0.20	−0.86***
	(0.17)	(0.12)	(0.15)
ln gross domestic product of country	1.20***	0.66***	1.60***
	(0.34)	(0.13)	(0.17)
Constant	−45.00***	−52.00***	−42.00***
	(13.00)	(16.00)	(4.40)
Observations	2,964	1,596	4,059
R^2	0.34	0.37	0.34
Adjusted R^2	0.33	0.37	0.34
ρ	0.94	0.88	1.00
Inverse Mills Ratio	4.10* (2.20)	3.50* (2.10)	4.50*** (0.81)
Note:			*p<0.1; **p<0.05; ***p<0.01



(a) ln gross domestic product 10th percentile

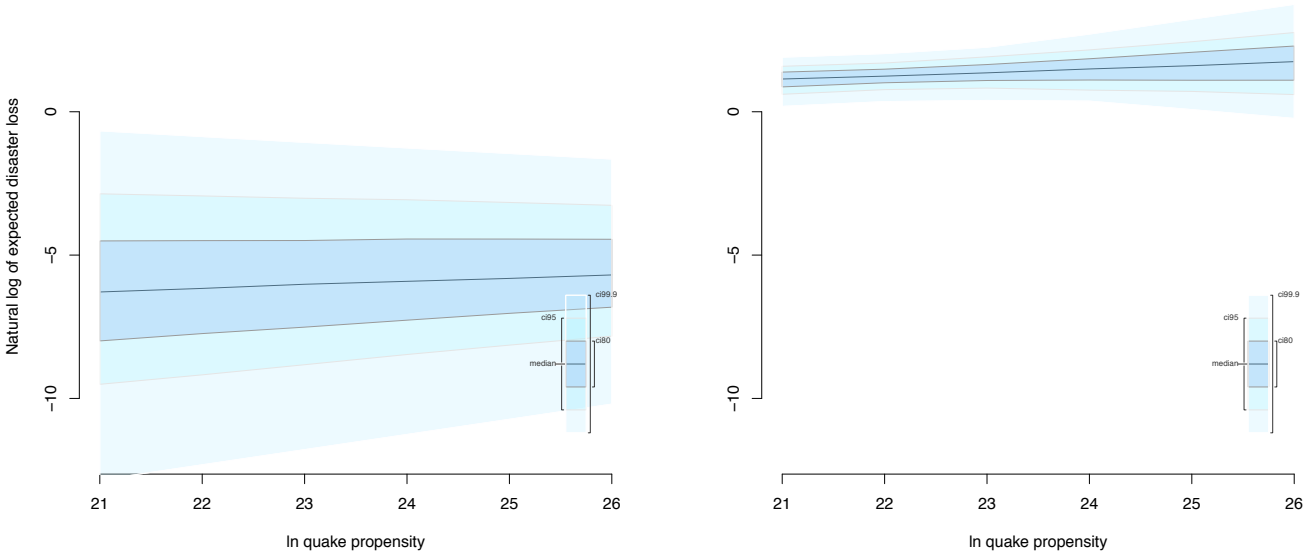
(b) ln gross domestic product 90th percentile

Figure 2: Effect of quake propensity on selection probability

holding disaster propensity constant, it is possible for wealthy individuals to invest extra resources in constructing disaster-proof buildings. Such efforts could effectively prevent any damages from weaker earthquakes and result in fewer number of quakes reported to Munich Re. Similar scenario is likely for floods, as damages from flood can also be prevented or mitigated by investing in flood-proof constructions.

Unlike wealthier individuals being associated with fewer disaster frequency, wealthier countries have more frequent observation of disasters. This is contrary to earlier studies where they found no correlation between the level of development and exposure to natural hazards [Kahn, 2005, Strömberg, 2007]. To investigate, we adapt Zelig to sample selection models and plot the effect of quake propensity on selection probability for countries in the 10th percentile vs. 90th percentile of GDP (Figure 2) [King et al., 2000]. As expected, among the countries in similar GDP quantile, countries with higher quake propensity are more likely to have a disaster observed in a country-year. We also see a clear contrast in selection probability by GDP – the countries with low GDP (Figure 2(a)) are less likely to have a disaster observation in the dataset than the countries with high GDP (Figure 2(b)) for all levels of quake propensity.

Various scenarios are possible. One is that frequent disasters are welfare improving. Some researches indeed suggest a positive economic impact of disaster, which we will discuss in the following section. Alternatively, the association could be an indication of selection bias where disasters from de-



(a) ln gross domestic product 10th percentile

(b) ln gross domestic product 90th percentile

Figure 3: Effect of quake propensity on log losses

veloping countries are more likely to be missing from Munich Re's dataset [Kron et al., 2012, Wirtz et al., 2014]. Identifying the validity of each scenario will be out of scope of this paper. Instead, we claim that the association between various covariates and the probability of observation shows that the original sample is not randomly selected. Thus, correcting the bias is crucial for our main analysis, the impact of disaster propensity on economic losses.

4.2 Economic loss

Table 2 shows the effect of disaster propensity, disaster hazard magnitude, GDP, and GDP per capita on the economic loss from earthquakes, cyclones, and floods. Not surprisingly, the higher the magnitude of a disaster, the larger the economic loss, as the coefficients of ln quake hazard magnitude are positive for all three disaster types (0.49,1.36, and 0.58, respectively).

Economic loss decreases with greater GDP per capita for floods (-0.86) is significant, but the effect is not signifcant for the other two disaster types. The result is fairly consistent with the estimates of [Neumayer et al., 2014], where it finds ln per capita income to be negatively associated with economic loss only for the 0.75 quantile for earthquakes, and 0.25, 0.5, and 0.75 quantiles for floods. Thus, the association between GDP per capita and reduced economic damage for floods is present regardless of selection bias.

Again, similar to the estimation of [Neumayer et al., 2014], economic loss increases with GDP for all three disaster types (1.64, 1.64, and 1.64). In [Neumayer et al., 2014], larger economic loss is associated with greater GDP for all quantiles of earthquakes, cyclones and floods except the 0.25 quantile of earthquake. The result is also in agreement with the stream of observations that richer countries also have more structures and wealth in hazardous areas, so damages tend to be higher [Raschky, 2008, Kousky, 2014].

In general, empirical evidences on the relationship between individuals' and countries' wealth and disaster damage have mixed results. Because our analysis is not designed to establish a causal claim between wealth and eco-

nomic loss, we recognize that various factors could be playing roles simultaneously to generate the result we see. Moreover, we observe that there is a higher level of uncertainty in economic losses reported for poor countries. Customizing Zelig to use the heckit variance-covariance, we plot the effect of quake propensity on economic loss for countries in the 10th percentile (Figure 3 (a)) vs. 90th percentile (Figure 3 (b)) of GDP [Toomet and Henningsen, 2008, King et al., 2000]. The poor countries tend to have significantly wider confidence interval compared to the rich countries. Since we cannot identify what is driving the wide differences in variances between rich vs. poor countries, we take it as a sign of limitation in using GDP and GDP per capita as a welfare measure when estimating their role in disaster mitigation and prevention.

4.3 Discussion: propensity and economic loss

Our primary relationship of interest is the effect of disaster propensity on economic loss. After correcting selection bias, disaster propensity is not negatively associated with economic loss. Instead, countries with higher disaster propensity have greater economic losses from cyclones (1.05) and floods (0.45), and although earthquake propensity does not turn out to be a significant predictor of economic loss, its direction is also positive (0.12).

One implication of our result is that correcting selection bias matters. In theory, selection bias threatens the internal validity of the result. In the analysis of [Neumayer et al., 2014], selection bias is obviously present and

significant. If we only select the country-years in which disasters are observed, we will not get a realistic estimate of the impact of disaster propensity, since the variable is correlated with selection. Fortunately, we are able to access the missing information about GDP and GDP per capita for observation-absent years, which is not always possible for other situations with selection bias. Furthermore, our bias corrected measure changes the direction of the effect, demonstrating that we should properly reflect the sample selection process of the data to strengthen the validity of results.

As we can see from Table 2, countries with higher disaster propensity tend to lose more than countries with lower disaster propensity when struck by the same level of disaster. Although we are cautious about making any causal claims, such result could indicate that disaster propensity is a measure of a country's vulnerability. We hypothesize two possible explanatory mechanisms.

First, frequent and severe disasters might have a long-lasting impact on the country's disaster prevention and mitigation systems, which leads to greater vulnerability towards economic loss for the next coming disaster. For example, destruction of infrastructures could take time to recover, which would contribute to greater economic losses if the country encounters another disaster without sufficiently recovering. Indeed, reconstruction of damaged or destroyed assets normally requires resources well beyond those available during the emergency stage, and reconstruction is often undertaken without vulnerability reduction [Bradshaw, 2003]. Thus, for countries that face fre-

quent and intense disasters, it is possible that vulnerability is reconstructed instead of being reduced.

Secondly, the negative association after correcting selection bias could be an indication that countries facing frequent disasters are not taking sufficient prevention and mitigation efforts to address their hazards properly. [Neumayer et al., 2014] argues that individuals and governments both face collective action problems, myopic behavior and asymmetric information, which all lead to insufficient disaster prevention and damage mitigation measures. Thus, the negative association between disaster propensity and economic loss can be an evidence of such disincentives. Similar evidence is present in [Schumacher and Strobl, 2011] – among wealthy countries, those with high disaster hazards are likely to have greater economic damage. Although the effect is only present among developed countries, the evidence also suggests that the level of country's mitigation and prevention effort might not be on par with what is necessary to prevent loss for countries that face frequent and intense disasters. Still, further research is necessary to understand the direction and the interaction of the impact of disaster propensity on the economy of natural disaster.

5 Conclusions and further works

What characteristics of a country suggest the likelihood of greater economic loss from a natural disaster? Are countries with higher disaster hazards more

susceptible to greater economic loss? To answer these questions, we investigate the relationship between disaster propensity, wealth, and economic loss from a panel data collected by [Neumayer et al., 2014]. We identify that the original data is subject to selection bias, apply Heckman correction, and find that countries that experience more frequent disasters tend to suffer from greater economic damage, holding everything else equal. Our result shows how simply addressing selection bias can change the direction and magnitude of the role of disaster propensity in economic losses. Still, we acknowledge that it is difficult to have a definitive explanation for the phenomenon, given the inherent limitations in estimating disaster losses and the macroeconomic nature of the question that requires addressing simultaneous time and geographical endogeneity. We conclude that when dealing with such complexities, statistical analysis should be applied with caution, and our analysis is one illustrative warning.

For future work, we suggest a more sophisticated construction of the latent propensity variable. We use the same propensity variable defined by [Neumayer et al., 2014] for our analysis, but the construction of variable bears an endogeneity problem, as we point out in Section 2.2. Although such problem is mitigated for pooled data, identifying an alternative propensity measure that is innately free from endogeneity concern could contribute to a more robust finding. Also, application of quantile selection models can help investigate the subtle differences in relationship by quantiles. Quantile regression enables a more detailed understanding of the impact of explanatory

variables than OLS, since we can estimate different effects of the explanatory variables at different points of the conditional disaster damage distribution. [d'Haultfoeuille et al., 2014] has proposed a novel methodology that allows quantile regression for models subject to endogenous selection. Taking advantage of the advance in methodological tools, we can make more nuanced claims about the relationship of interest.

In the context of climate change, understanding the role of institutional, political, and other national conditions in determining disaster impacts is imperative. We believe addressing threats to validity is especially crucial in disaster economy literature for establishing a meaningful empirical evidence.

References

[WB, 2014] (2014). World development indicators 2014. *Washington, DC: World Bank*.

[Bradshaw, 2003] Bradshaw, S. (2003). Handbook for estimating the socio-economic and environmental effects of disasters.

[d'Haultfoeuille et al., 2014] d'Haultfoeuille, X., Maurel, A., and Zhang, Y. (2014). Extremal quantile regressions for selection models and the black-white wage gap. Technical report, National Bureau of Economic Research.

[Fischer and Knutti, 2015] Fischer, E. M. and Knutti, R. (2015). Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nature Clim. Change*, advance online publication:–.

[Heckman, 1979] Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161.

[IFRC, 2014] IFRC (2014). World disaster report 2014: Focus on culture and risk. In *World Disaster Report 2014: Focus on culture and risk*. International Federation of Red Cross and Red Crescent Societies (IFRC).

[Kahn, 2005] 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.

[Kellenberg and Mobarak, 2008] Kellenberg, D. K. and Mobarak, A. M. (2008). Does rising income increase or decrease damage risk from natural disasters? *Journal of Urban Economics*, 63(3):788 – 802.

[King et al., 2000] King, G., Tomz, M., and Wittenberg, J. (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American journal of political science*, pages 347–361.

[Kousky, 2014] Kousky, C. (2014). Informing climate adaptation: A review of the economic costs of natural disasters. *Energy Economics*, 46(0):576 – 592.

[Kron et al., 2012] Kron, W., Steuer, M., Löw, P., and Wirtz, A. (2012). How to deal properly with a natural catastrophe database? analysis of flood losses. *Natural Hazards and Earth System Science*, 12(3):535–550.

[Neumayer et al., 2014] Neumayer, E., Plümper, T., and Barthel, F. (2014). The political economy of natural disaster damage. *Global Environmental Change*, 24:8–19.

[Raschky, 2008] Raschky, P. A. (2008). Institutions and the losses from natural disasters. *Natural Hazards and Earth System Science*, 8(4):627–634.

[Sartori, 2003] Sartori, A. E. (2003). An estimator for some binary-outcome selection models without exclusion restrictions. *Political Analysis*, 11(2):111–138.

[Schumacher and Strobl, 2011] Schumacher, I. and Strobl, E. (2011). Economic development and losses due to natural disasters: The role of hazard exposure. *Ecological Economics*, 72(0):97 – 105.

[Strömberg, 2007] Strömberg, D. (2007). Natural disasters, economic development, and humanitarian aid. *The Journal of Economic Perspectives*, pages 199–222.

[Toomet and Henningsen, 2008] Toomet, O. and Henningsen, A. (2008). Sample selection models in R: Package sampleSelection. *Journal of statistical software*, 27(7):1–23.

[Toya and Skidmore, 2007] Toya, H. and Skidmore, M. (2007). Economic development and the impacts of natural disasters. *Economics Letters*, 94(1):20–25.

[Tucker, 2013] Tucker, B. E. (2013). Reducing earthquake risk. *Science*, 341(6150):1070–1072.

[Willmott and Matsuura, 2011] Willmott, C. and Matsuura, K. (2011). Willmott, matsuura and collaborators' global climate resource pages.

[Wirtz et al., 2014] Wirtz, A., Kron, W., Löw, P., and Steuer, M. (2014). The need for data: natural disasters and the challenges of database management. *Natural hazards*, 70(1):135–157.

A Panel

Figure 4 and 5 show that countries face varying levels of disaster hazards and economic losses. Some countries have sparse observations compared to others, and it is not clear whether those difference come from the differences in disaster vulnerability, or selection bias.

B Confidence interval plots

Figure 6 and 8 show the effect of disaster propensity on selection probability by GDP for cyclones (Figure 6) and floods (Figure 8). As expected, we see

that among the countries with similar GDP, higher disaster propensity is associated with higher likelihood of observation in a country-year. Richer countries have higher probability of observing a disaster than poor countries for both disaster types.

Figure 7 and 9 show the effect of disaster propensity on economic losses by GDP for cyclones (Figure 7) and floods (Figure 9). Poor countries tend to have significantly wider confidence intervals for expected loss compared to rich countries for both disaster types.

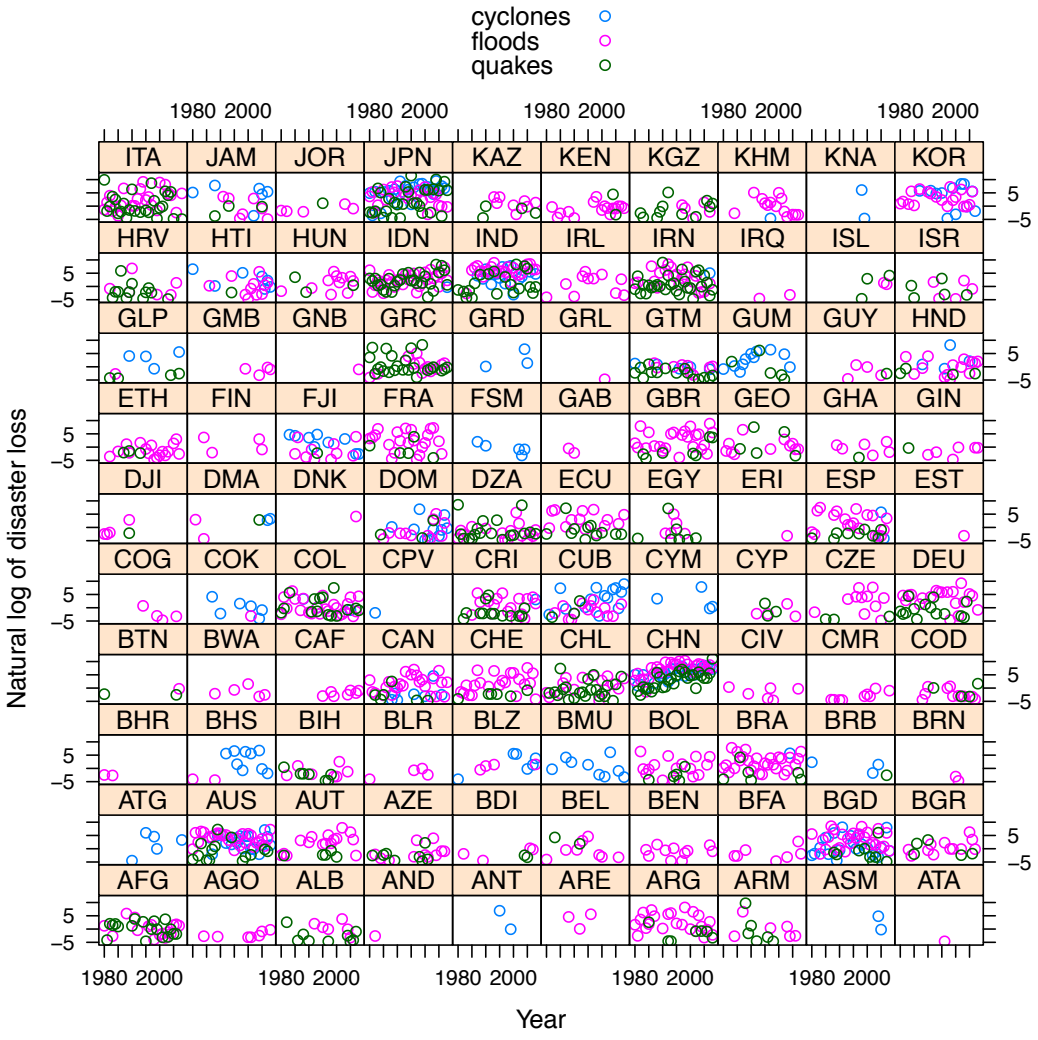


Figure 4: Log losses by disaster type per country for period 1980-2008

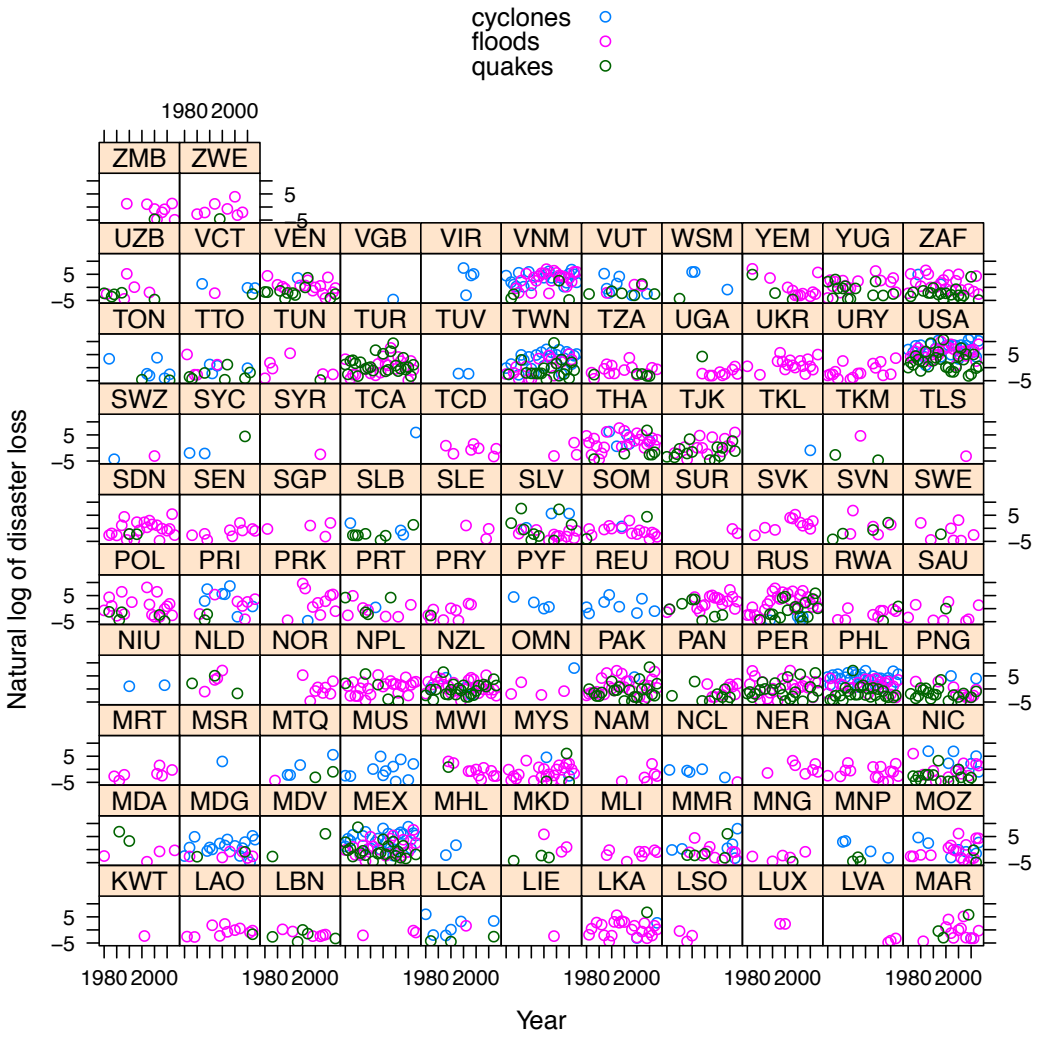
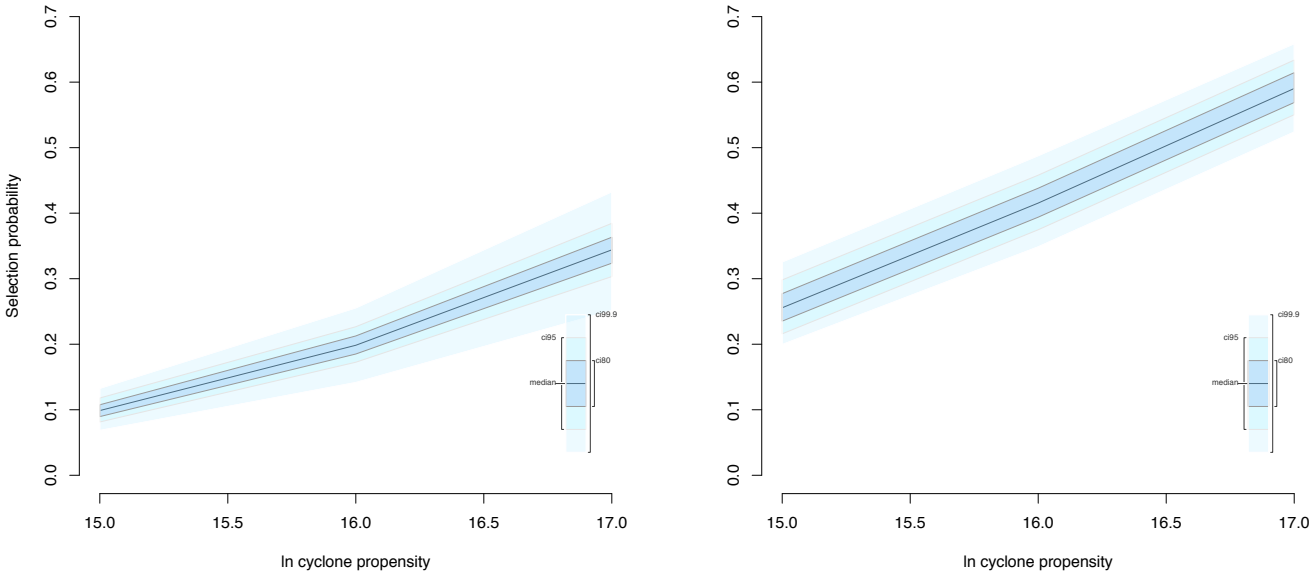


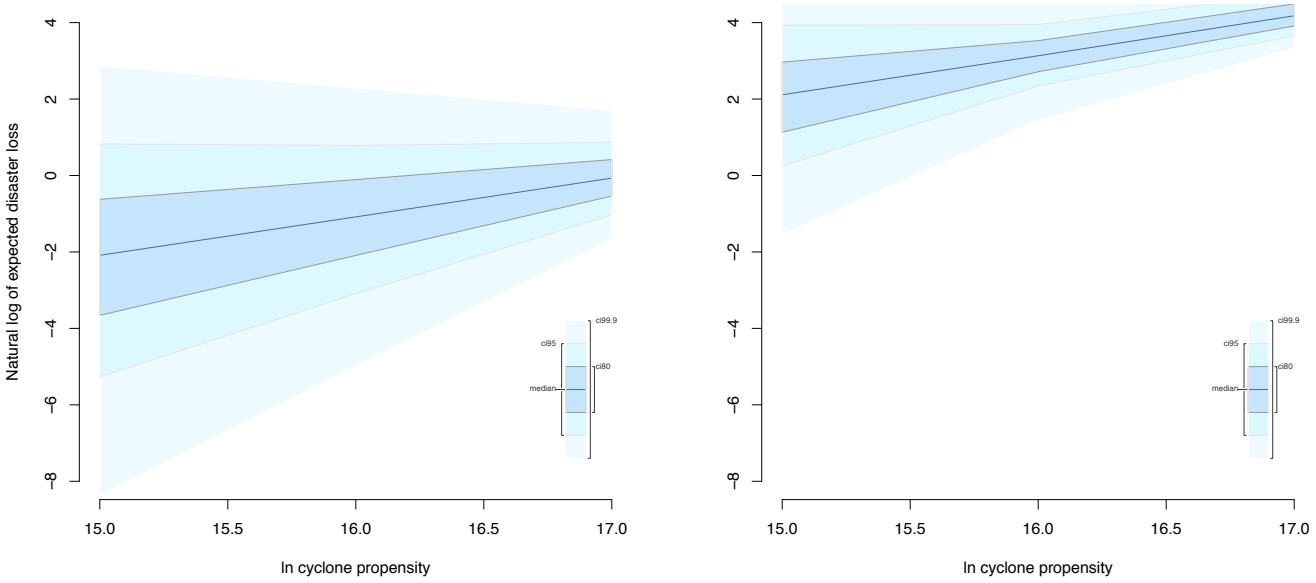
Figure 5: Log losses by disaster type per country for period 1980-2008



(a) ln gross domestic product 10th percentile

(b) ln gross domestic product 90th percentile

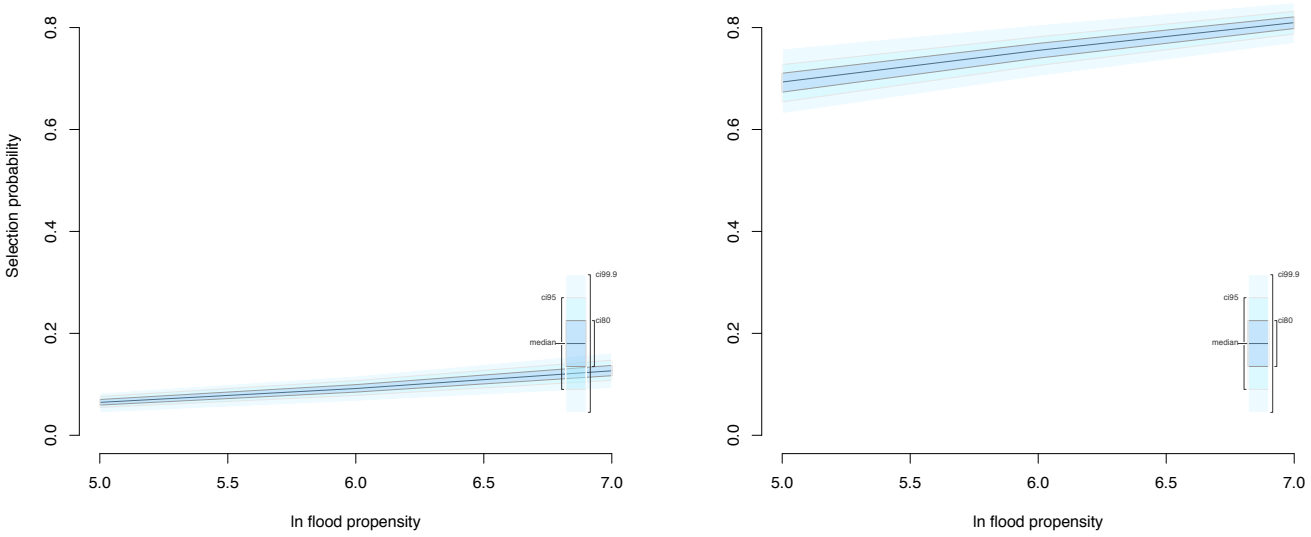
Figure 6: Effect of cyclone propensity on selection probability



(a) ln gross domestic product 10th percentile

(b) ln gross domestic product 90th percentile

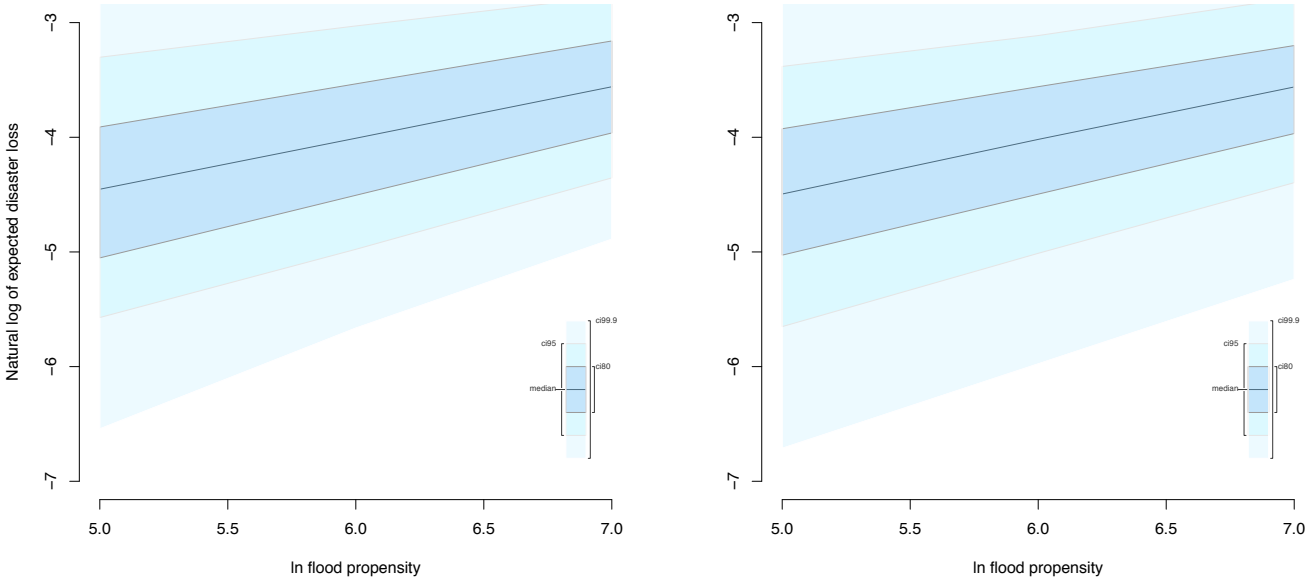
Figure 7: Effect of cyclone propensity on log losses



(a) ln gross domestic product 10th percentile

(b) ln gross domestic product 90th percentile

Figure 8: Effect of flood propensity on selection probability



(a) ln gross domestic product 10th percentile

(b) ln gross domestic product 90th percentile

Figure 9: Effect of flood propensity on log losses