Supplementary material for "MORE INEQUALITY, MORE KILLINGS: THE MAOIST INSURGENCY IN NEPAL"

1. *Measurement of inequality:*

   The Gini index for each village according to Deaton's (2000, p. 139) measure is:

$$GINI = \frac{1}{\mu N(N-1)} \sum_{i>j} \sum_{j} | y_i - y_j | , \qquad [2a]$$

where y_l denotes per capita consumption expenditure of household *l* in a given village. μ is average expenditure, *N* is sample size, and $| y_i - y_j |$ is the absolute deviation of expenditure between two households. A closely related formulation of the Gini index is:

$$GINI = \frac{N+1}{N-1} - \frac{2}{N(N-1)\mu} \sum_{i=1}^{N} \rho_i y_i \qquad [2b]$$

where ρ_i is the rank of household *i* in the *y*-distribution, arranged so that the richest person is ranked 1. We use [2b] to compute our village-level Gini.

Esteban and Ray's (1999) concept of polarization is designed to measure rivalry among homogeneous groups (the Gini is a special case of their polarization index, were the group size is one). It is therefore a measure of horizontal or cross-group inequality. Axiomatically, their measure purports to satisfy three conditions: First, the joining of two neighboring probability masses into one mass exacerbates polarization in the presence of another separately identifiable point mass. Second, given three point masses (or partitions), moving a point mass away from the center towards an extreme value, however small the move, increases polarization. Third, given two point masses (or partitions), breaking the more central point mass equally into two and distributing them at two opposite, more extreme, points increases polarization. Satisfying these conditions yields a measure that may not necessarily be correlated with the Gini since it seeks to

measure potential hostility or antagonism among groups, and therefore captures a *horizontal* dimension of inequality different than the Gini. In our case, the Esteban-Ray economic polarization measure maps the distribution of consumption spending by households in a village into a value for the village. The higher this value, the greater is the degree of polarization within the village. Esteban and Ray's (1999) polarization index for a village is given by:

$$POLARIZATION(\alpha) = K \sum_{i=1}^{L_i} \sum_{j=1}^{N_j} \pi_i^{1+\alpha} \pi_j \mid y_i - y_j \mid, \qquad [3]$$

where $| y_i - y_j |$ is the size of absolute difference in the consumption expenditure of households i and j, π_k is the k^{th} household's proportional weight[1] and L_k is the number of households sampled from k^{th} village. K is a positive constant. α measures the intensity of group identification, termed the "degree of polarization sensitivity" by Esteban and Ray, and shown to range between 0 and 1.6. The Gini index is the special case of [3] with α=0.[2] The larger is α, the greater is the departure of the inequality measure from the Gini. Using the kernel estimation method of Duclos, Esteban and Ray (2004), we construct an Esteban-Ray polarization measure at the village level based on α=1. The measure is termed POLARIZATION(α=1).

Four other measures of horizontal inequality explicitly based on partitioning groups according to religion, language and caste (they partition society into those that are upwardly mobile from those that aren't). We first use the Cederman et al. (2010) measure:

$$\text{HI}(g) = (\ln(y_g / \bar{y}))^2,$$

where y_g is the per capita income of group g (in the village), and $\bar{y}$ is average per capita income over the sample of villages. We construct this measure for three traditionally underprivileged groups g: (i) lower caste (non- Brahmin or non-Chhetri), (ii) non-Nepali speaking, and (iii) non-

[1] Since the census sample includes only one in eight households, the household's proportional weight uses information about population size of each village and the size of the census-sampled households.

[2] Montalvo and Reynal-Querol (2005) show that $GINI = \sum_{i=1}^{L_i} \sum_{j=1}^{N_j} \pi_i \pi_j \mid y_i - y_j \mid$,

Hindu, respectively termed on HINEQUALITY(C), HINEQUALITY(L), and HINEQUALITY(C1). The fourth HI measure, based on Mancini (2005), partitions based on ethnicity:

$$\text{GCOV} = \frac{1}{\bar{y}}\left(\sum_{r}^{R} p_r\,(\bar{y}_r - \bar{y})^2\right)^{\frac{1}{2}}$$

where R is the number of ethnic groups in a village,[3] $\bar{y}$ is overall mean of income across all villages, $\bar{y}_r$ is the mean income for ethnic group r, and p_r is group r's population share. This measure is termed on HINEQUALITY(C2).

Poverty measure

The poverty measure we use is the Foster-Greer-Thorbecke (FGT) poverty-gap index for the year 1995-96. It calculates the percentage of households in a village below the poverty line:

$$POVERTY = \frac{1}{n}\sum_{i \in L_p}^{q} \left(\frac{z - y_i}{z}\right)^{\gamma}, \qquad [4]$$

where z defines a household's poverty expenditure threshold,[4] y_i is household i's expenditure, n is the number of households, and L_p is the set of households below the poverty line (the notation is simplified: n and L_p vary across villages). $\gamma > 0$ is a poverty aversion parameter: if $\gamma = 0$, [4] is simply the proportion of households below the poverty line, or the "headcount" index. When $\gamma = 1$, [4] is the average poverty-gap index, that is, the average shortfall of household expenditure from the poverty line. The poverty measures for different values of γ are highly correlated in our sample, and we use the headcount measure ($\gamma = 0$) in the analysis.

[3] Seven ethnicities are used to partition the population of each village: Brahmin/Chhetri, Tamagurali (Tamang, Magar, Gurung, Rai, and Limbu), Dakasa (Damai, Kami, Sarki), Newar, Muslim, Terai Caste (lower "untouchable" caste from Terai), and Other Caste.

[4] We adopted the poverty line definition used by Nepal's Central Bureau of Statistics of 4404 Nepali rupees at 1995/96 constant price (CBS 2005).

The simple correlations among the inequality and poverty variables indicate that many of them measure distinct and different dimensions of inequality.

(*N*=3857)

	GINI	POL	POV	HI(L)	HI(R)	HI(C1)	HI(C2)
GINI	1.00						
POLARIZATION	0.47	1.00					
POVERTY	−0.28	0.008	1.00				
HINEQUALITY(L)	−0.38	−0.24	0.59	1.00			
HINEQUALITY(R)	−0.19	−0.09	0.31	0.55	1.00		
HINEQUALITY(C1)	−0.11	0.002	0.21	0.31	0.48	1.00	
HINEQUALITY(C2)	−0.24	−0.18	0.36	0.77	0.52	0.34	1.00

Districts:

ID	District
1	Taplejung
2	Panchthar
3	Ilam
4	Jhapa
5	Morang
6	Sunsari
7	Dhankuta
8	Terhathum
9	Sankhuwasabha
10	Bhojpur
11	Solukhumbu
12	Okhaldhunga
13	Khotang
14	Udayapur
15	Saptari
16	Siraha
17	Dhanusa
18	Mahottari
19	Sarlahi
20	Sindhuli
21	Ramechhap
22	Dolakha
23	Sindhupalchok
24	Kavre
25	Lalitpur
26	Bhaktapur
27	Kathmandu
28	Nuwakot
29	Rasuwa
30	Dhading
31	Makwanpur
32	Rautahat
33	Bara
34	Parsa
35	Chitwan
36	Gorkha
37	Lamjung
38	Tanahu
39	Syangja
40	Kaski
41	Manang
42	Mustang
43	Myagdi
44	Parbat
45	Baglung
46	Gulmi
47	Palpa
48	Nawalparasi
49	Rupandehi
50	Kapilbastu
51	Arghakhanchi
52	Pyuthan
53	Rolpa
54	Rukum
56	Dang
57	Banke
58	Bardiya
59	Surkhet
60	Dailekh
61	Jajarkot
62	Dolpa
63	Jumla
65	Mugu
66	Humla
67	Bajura
68	Bajhang
69	Achham
70	Doti
71	Kailali
72	Kanchanpur
73	Dadeldhura
74	Baitadi
75	Darchula

Table A1: First Stage regressions for endogenous variables

		Village Level Instruments						District level instruments					
		GINI	POL (α=1)	HI(L)	HI(R)	HI(C1)	HI(C2)	GINI	POL (α=1)	HI(L)	HI(R)	HI(C1)	HI(C2)
Excluded Instruments	%WithAGLAND	-0.0502***	-0.0151***	-0.0953**	-0.0897**	-0.144***	-0.198***	-	-	-	-	-	-
		[0.00975]	[0.00294]	[0.0394]	[0.0415]	[0.0496]	[0.0557]	-	-	-	-	-	-
	%WithFEMALELAND	0.0533***	0.00777***	-0.0634*	0.0117	0.0282	-0.0266	-	-	-	-	-	-
		[0.0106]	[0.00290]	[0.0328]	[0.0421]	[0.0515]	[0.0317]	-	-	-	-	-	-
	BIGLIVESTOCK	-0.00329	-0.00171*	0.0139	0.0311*	0.0287	0.00241	-	-	-	-	-	-
		[0.00300]	[0.000926]	[0.0103]	[0.0180]	[0.0237]	[0.00749]	-	-	-	-	-	-
	RAIN Variance	-	-	-	-	-	-	0.00243	-0.000857	-0.0458***	-0.02	-0.0281	-0.0342***
		-	-	-	-	-	-	[0.00301]	[0.00102]	[0.0102]	[0.0136]	[0.0237]	[0.00901]
	RAIN Precision	-	-	-	-	-	-	6.20e-05***	8.87e-06***	0.000145***	0.000158***	0.000215***	0.000112***
		-	-	-	-	-	-	[8.31e-06]	[2.75e-06]	[3.10e-05]	[3.65e-05]	[5.41e-05]	[1.92e-05]
Included Exogenous variables	POVERTY	-0.0870***	0.0019	1.021***	0.796***	0.675***	0.642***	-0.0724***	0.000814	0.624***	0.323***	0.232***	0.305***
		[0.00772]	[0.00268]	[0.0615]	[0.0507]	[0.0608]	[0.0577]	[0.00381]	[0.00126]	[0.0182]	[0.0235]	[0.0251]	[0.0181]
	EDUCATION	0.000517	0.000164	-0.00367	-0.0033	-0.000515	0.00392	0.000742	0.000299	-0.00539**	-0.000306	0.00143	0.00304
		[0.000968]	[0.000295]	[0.00316]	[0.00567]	[0.00606]	[0.00372]	[0.000658]	[0.000210]	[0.00262]	[0.00333]	[0.00375]	[0.00238]
	EMPLOYMENT	5.84E-05	0.000275**	-0.0113***	-0.0107***	-0.0101***	-0.00493***	-0.000354	0.000266**	-0.0165***	-0.0145***	-0.0121***	-0.00780***
		[0.000357]	[0.000126]	[0.00126]	[0.00226]	[0.00264]	[0.00110]	[0.000372]	[0.000122]	[0.00138]	[0.00207]	[0.00250]	[0.00107]
	FARMER	-0.0290***	-0.00706***	-0.00946	-0.0337	0.0569	-0.0294	-0.0449***	-0.0130***	0.148***	0.152***	0.117***	-0.00116
		[0.00623]	[0.00245]	[0.0325]	[0.0441]	[0.0471]	[0.0317]	[0.00530]	[0.00167]	[0.0212]	[0.0268]	[0.0306]	[0.0211]
	ETHNICITY	0.00323	0.000623	0.0448***	-0.0409*	-0.0658**	0.000348	0.0139***	-0.00264***	0.0633***	-0.0928***	-0.153***	0.0158***
		[0.00293]	[0.000779]	[0.0143]	[0.0231]	[0.0279]	[0.0150]	[0.00181]	[0.000566]	[0.00700]	[0.00999]	[0.0106]	[0.00565]
	RURAL	0.0412***	4.81E-06	-0.160***	-0.100***	-0.0674**	-0.141***	0.0110*	-0.00390**	-0.169***	-0.0591**	-0.00828	-0.159***
		[0.00751]	[0.00216]	[0.0304]	[0.0308]	[0.0307]	[0.0414]	[0.00660]	[0.00181]	[0.0300]	[0.0256]	[0.0261]	[0.0404]
	ln(POPULATION)	0.0453***	0.00287***	0.0035	0.0227***	0.0190*	0.0044	0.0373***	0.00283***	-0.0189***	0.0242***	0.0333***	-0.00598
		[0.00258]	[0.000534]	[0.00614]	[0.00789]	[0.00979]	[0.00826]	[0.00143]	[0.000354]	[0.00424]	[0.00558]	[0.00637]	[0.00411]
	N	3857	3857	3857	3857	3857	3857	3857	3857	3857	3857	3857	3857
	R^2	0.528	0.075	0.47	0.2	0.113	0.269	0.439	0.099	0.457	0.137	0.113	0.179
	District Fixed Effects	YES	YES	YES	YES	YES	YES	NO	NO	NO	NO	NO	NO
	First stage *F*	15.6	12.81	2.93	2.91	4.53	4.34	27.97	6.33	25.96	12	10.26	27.52

Notes:

1. Standard errors in parentheses. ***, **, and * indicate statistical significance at 1%, 5% and 10%, respectively.
2. First stage *F* is used to test the joint null that %WithAGLAND = %WithFEMALELAND = BIGLIVESTOCK =0 in FE models and RAINVariance=RAINPrecision=0 in non-FE models..

Table A2: Auxiliary regression with Instruments

Dependent variable: Ln(KILLINGS)				
%WithAGLAND	−1.406***	−1.753***	−1.079**	−1.480**
	[0.491]	[0.595]	[0.482]	[0.582]
%WithFEMALELAND	2.652***	2.155***	2.499***	2.042***
	[0.702]	[0.712]	[0.691]	[0.703]
BIGLIVESTOCK	−0.27	−0.294	−0.256	−0.283
	[0.268]	[0.269]	[0.265]	[0.268]
GINI	−2.068	−2.001		
	[1.794]	[1.804]		
POLARIZATION (α=1)			9.061**	9.049**
			[3.941]	[3.960]
POVERTY	−0.216	−0.261	−0.053	−0.125
	[0.525]	[0.601]	[0.490]	[0.563]
EDUCATION		−0.0342		−0.0398
		[0.0663]		[0.0667]
EMPLOYMENT		−0.0610*		−0.0608*
		[0.0322]		[0.0321]
FARMER		0.678		0.809
		[0.634]		[0.640]
ETHNICITY		0.668***		0.662***
		[0.235]		[0.239]
RURAL		0.00074		−0.00766
		[0.314]		[0.306]
ln(POPULATION)	1.177***	1.162***	1.046***	1.039***
	[0.122]	[0.139]	[0.108]	[0.122]
N	3857	3857	3857	3857
k	82	87	82	87
FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.105	0.107	0.105	0.107