**Supporting Information**

***Methods***

*Socioeconomic data cleaning*

Prior to aggregation to the commune level, village data were checked for missing values. In some cases, villages had data for a subset of years but were missing data for other years. If the missing data were at the start of the study period or the end of the study period it was assumed that the village was either an old or a new village. Villages can be merged with larger villages, or two sub-villages, or “*Kroms*”, can be split into two distinct villages over time for administrative purposes. In these cases, the rows (years) with missing data were deleted, but the years with data were retained as these represent villages that existed in that year. If the missing data were in the middle of the study period (for more than one year), or if data for that village only exists for one or several years in the middle of the study period, then the data were assumed to be incomplete and the village was deleted. If the village had data for all years except one, then the missing values were estimated using linear interpolation. If the village existed in all years, but was missing data from multiple years, the village was deleted. If an entire commune was missing in some years, the commune was deleted. The above cleaning process removed 312 villages (total number of villages = 84,195), or 0.37% of the data. Data were then split into individual years, and the final village-level data were aggregated to the commune level using the operations defined below in Table Sx.

**Table Sx. Mathematical operations used to aggregate socioeconomic variables from the village to the commune level.**

|  |  |
| --- | --- |
| Variable | Operation |
| Total population | Sum |
| Number of families | Sum |
| Number of males aged 18-64 | Sum |
| Number of females aged 18-64 | Sum |
| Number of people aged over 61 | Sum |
| Total number of indigenous people | Sum |
| Number of families whose main occupation is farming | Sum |
| Number of land conflict cases | Sum |
| Number of in-migrants | Sum |
| Number of out-migrants | Sum |
| Number of criminal cases | Sum |
| Proportion of population that is indigenous | Mean |
| Proportion of females aged 6-24 in full time education | Mean |
| Proportion of males aged 6-24 in full time education | Mean |
| Proportion of females aged 15-45 who are illiterate | Mean |
| Proportion of males aged 15-45 who are illiterate | Mean |
| Proportion of families whose main occupation is farming | Mean |
| Proportion of people who are primarily employed in the primary sector | Mean |
| Proportion of people who are primarily employed in the secondary sector | Mean |
| Proportion of people who are primarily employed in the tertiary sector | Mean |
| Proportion of people who are primarily employed in the quaternary sector | Mean |
| Proportion of families who have less than 1ha of farmland | Mean |
| Proportion of families who have buffalo | Mean |
| Proportion of families who have pigs | Mean |
| Proportion of families who have access to waste collection | Mean |
| Number of infant (<6mo) mortality cases | Mean |
| Number of child (<5 years old) mortality cases | Mean |
| Distance to the nearest school | Median |
| Distance to the Commune Office | Median |
| Distance to the nearest health centre | Median |

After aggregation, each variable was checked for obvious errors or unlikely outliers via plotting of histograms and trends. Plots were done at the province level first, to identify any communes within a province that had particularly unusual values or trends. If unusual values or trends were identified the commune was investigated in more detail.