Understanding Hsiang and Sekar's analysis

Hsiang and Sekar (H&S) have written a rebuttal to my critique with Bob Burn regarding their paper claiming that the legal sale led to an increase in poaching rates. In this rebuttal they claim that (1) simple plots of the data show a clear discontinuity in illegal killing of elephants between 2007 and 2008 and (2) the main reason for the discrepancy between their analysis and Burn et al (2011) is because we fitted smooth terms to describe the trend through time.

As a member of the MIKE-ETIS Technical Advisory Group I was asked if could look at their analysis and understand why the results differ so much. This report shows that they have ignored a key feature of the PIKE data and that it is this rather than anything to do with smoothing that is the reason why their analysis differs from ours, and all other analyses of the MIKE data.

In their response, H&S provide a simplified version of the main result from their paper which shows a discontinuity between 2007 and 2008 – see Figure 1A.

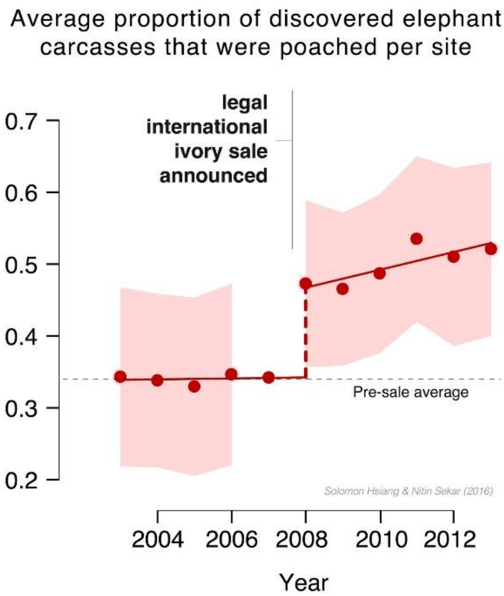

Figure 1A Hsiang and Sekar's main result showing a discontinuity.

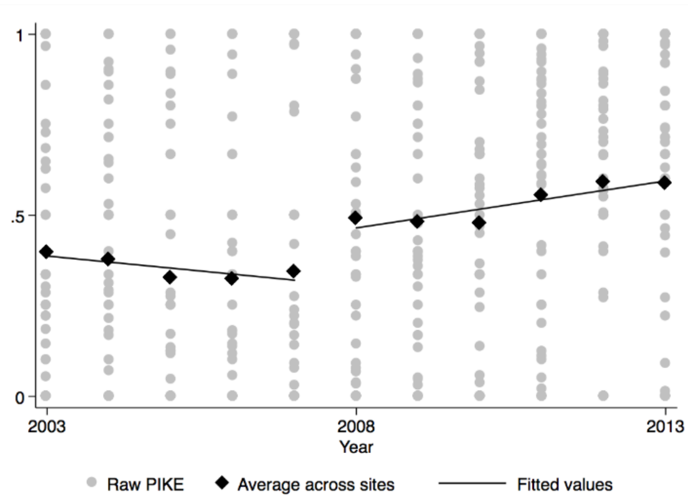

Figure 1B Hsiang and Sekar's graph showing PIKE for each site and the average for each site

My understanding is that the dots in this are from a model which calculates average PIKE having removed the site means, and the lines are a regression which allow for different trends both before and after 2008. They claim that the discontinuity can be seen by a very simple summary of the data shown in Figure 1B on the right. In this plot, the grey dots are the PIKE values for each site, note how they vary hugely between sites, and the black diamonds are the mean PIKE values for each year. You can see that the points in both figures look quite similar.

In their paper they also show the aggregate PIKE values for each year (Figure 2). Aggregate PIKE values are calculated as the sum of all the illegally killed carcasses that were found and reported from all sites divided by the sum of all of the carcasses that were found and reported across all sites. In their paper aggregate PIKE is the blue line in this figure when they have removed three data points[1].

[1] Two in 2009 from Kenya (SBR and TSV) and one in 2012 from Botswana (CHO)

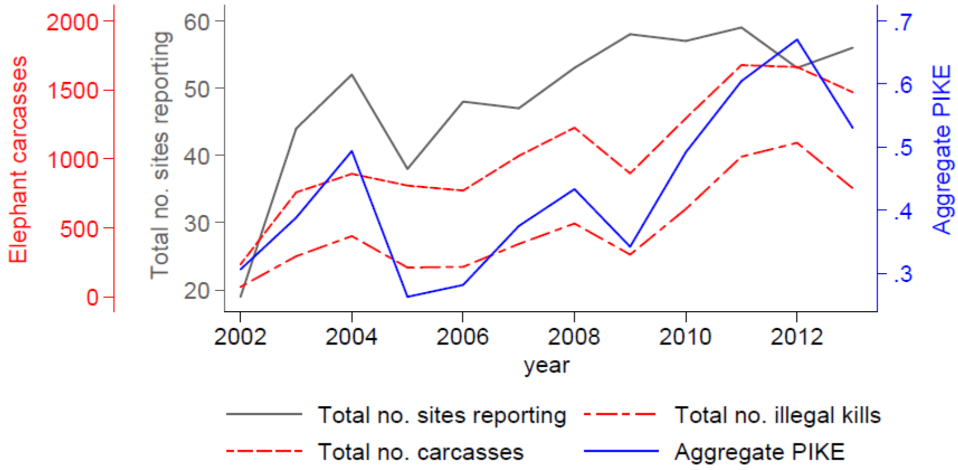

Figure 2: Hsiang and Sekar's graph showing aggregate PIKE

I have calculated the aggregate and average PIKE values using their data (Figure 3). Note that I have removed 2002 (as they have excluded them from their analyses) and six cases where no carcasses were found at all[2].

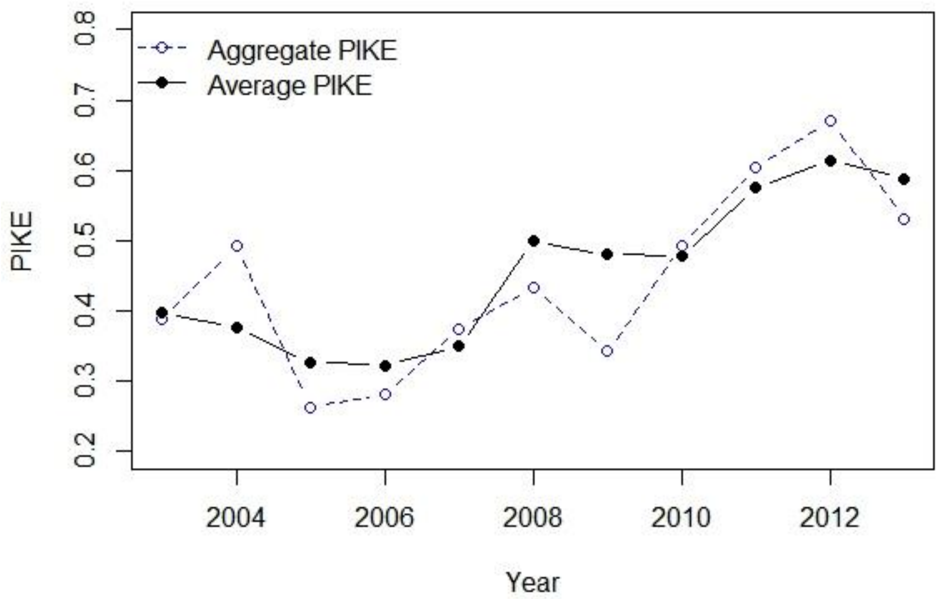

Figure 3: Plot of Hsiang and Sekar's average and aggregate PIKE removing observations that H&S remove and observations where no carcasses were found at all.

There are two points to make about this.

[2] because these results are indeterminate - although I note they have set PIKE to 0 for these cases. The fact that these have been removed doesn't make any difference to any of the results that I show but I don't feel comfortable using them and they will be ignored in most analyses anyway.

1. The first is that the aggregate value (blue dashed line) looks much more similar to other analyses of the MIKE data that do not show a sharp discontinuity between 2007 and 2008.

2. The two lines diverge in a number of places – in particular the period 2005 - 2009. Both are very simple summaries of the data so it is important to understand why these differ. This is, in fact, the key to understanding why Hsiang and Sekar's analysis is so different to all other analyses.

Let's review what PIKE is. It is the number of elephant carcasses that were found at a site that were illegally killed divided by the total number of elephant carcasses that were found:

$$PIKE = \frac{number\ of\ illegally\ killed\ carcasses\ found\ at\ the\ site}{total\ number\ of\ carcasses\ found\ at\ the\ site}$$

If the total number of carcasses found at each site was exactly the same, say 50, then, mathematically, the aggregate and average PIKE values would look exactly the same, irrespective of the variability in individual PIKE values. But that isn't the case here. For example, in 2012 there are 24 sites which each found no more than 10 carcasses and six sites where at least 100 carcasses were found at each site. Figure 4 below shows how the numbers differ within in each year and between years.

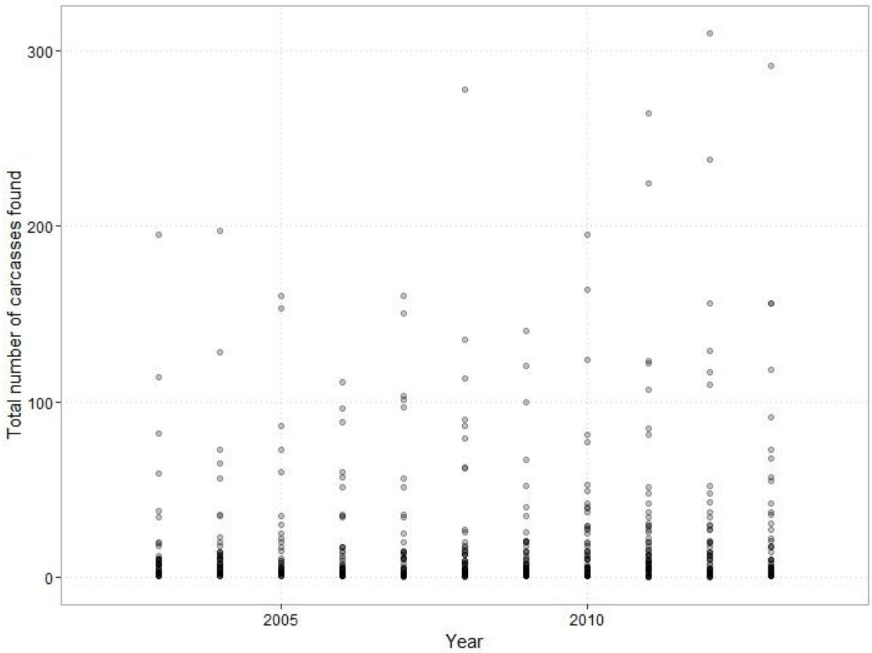

Figure 4: Total number of carcasses found at each site in each year. Dark areas indicate that there are lots of sites where the same number of carcasses were found.

This feature of the data has been completely ignored by Hsiang and Sekar in any discussion, or analysis.

Why does it matter and which is correct?

Let's think about what PIKE actually represents. PIKE is the proportion of carcasses that were found at a site that were illegally killed. But the number of carcasses found does not, in general, represent all of the elephant deaths at that site. For example, we know that the estimate of the number of elephants that were illegally killed from MIKE sites in Central Africa in 2012 was over 10,000 (Wittemyer et al 2013). But only 195 carcasses were found in total from these sites, of which 169 were illegally killed. So the carcasses that are found are a "sample[3]" of the total number of elephants that were killed. If a wildlife patrol went somewhere different they would find a different number of carcasses of which a different proportion were illegally killed. The value of PIKE that we get is an estimate of the true proportion of elephants that died that were illegally killed.

Consider the situation where one extra carcass is found at a site as illustrated in Table 1.

	Baseline			One additional carcass found						
	Total found	Illegally killed	PIKE	Total found (+1)	Illegally killed (+1)	PIKE	Change from baseline	Illegally killed (no change)	PIKE	Change from baseline
Site A	5	2	0.4	6	3	0.5	0.1	2	0.33	0.07
Site B	50	20	0.4	51	21	0.41	0.01	20	0.39	0.01

Table 1: Sensitivity of PIKE when one additional carcass is found differs with total number of carcasses.

If initially there were five carcasses found at a site of which two were illegally killed, PIKE is 0.4. If that extra carcass was also illegally killed, PIKE changes to 0.5 (an increase of 0.1). If the extra carcass was not illegally killed, then PIKE changes to 0.33 (a decrease of 0.07). In comparison if 50 carcasses are found of which 20 were illegally killed, PIKE is again 0.4. If one more carcass is found and it was illegally killed, PIKE is 0.41 (an increase of 0.01). If the carcass was not illegally killed, then PIKE is 0.39 (a decrease of 0.01).

In other words, when only a small number of carcasses are found at a site the estimate of PIKE is extremely sensitive to the addition of one extra carcass – this is not the case when the number of carcasses that have been found is large. More technically, we would say that the variability in our estimate of PIKE increases as the total number of carcasses that have been found decreases (everything else remaining equal).

It is important to account for this when calculating the average PIKE estimate. Specifically, you want to give more weight to sites where the proportion is calculated from a large number of carcasses than sites where the proportion is based on a small number of carcasses. You can do this by calculating a weighted average of the data. For proportions, a sensible weighting is to use the total number of carcasses found at a site – doing this gives you aggregate PIKE.

Any analysis of the data must also take account of this issue. In particular, a key assumption of the regression analysis is that all data points are equally precise estimates of whatever it is they are measuring. This is already violated to some extent when we have proportions but isn't a major issue if the number of carcasses is relatively similar. But here the differences are quite large – in fact this is illustrated by the fact that the weighted average or aggregate PIKE looks so different to the straight average of PIKE. Given the general approach that Hsiang and Sekar have taken one way to tackle this problem would be to do a weighted regression where the weights represent the total number of carcasses are found.

A second way to deal with proportions when you know the numerator and denominator is to fit a generalised linear model (GLM), assuming a Binomial distribution for the data. I cannot stress enough that this approach has been widely used in many field of application for at least 40 years and

[3] I have put sample in quotes because it is collected on wildlife patrols rather than using any kind of sampling strategy.

is completely uncontroversial. This approach deals automatically with the problem of varying number of total carcasses found and other issues which have previously been outlined – such as the fact that the data are constrained to be between 0 and 1.

A third approach would be to fit a GLM and assume a Poisson distribution for the number of illegally killed carcasses. Although Hsiang and Sekar claim to have done so their model is in fact wrong[4].

Comparison of modelling approaches

I have fitted all of these three models and compared it to that of Hsiang and Sekar. These models are very basic models[5] and ignore more complex issues such as the fact that some sites come from the same countries etc and also do not do any smoothing. These are frills (relatively)and your basic models will give a general idea of what is going on. The results of these model fitting exercises are shown in Figure 5.

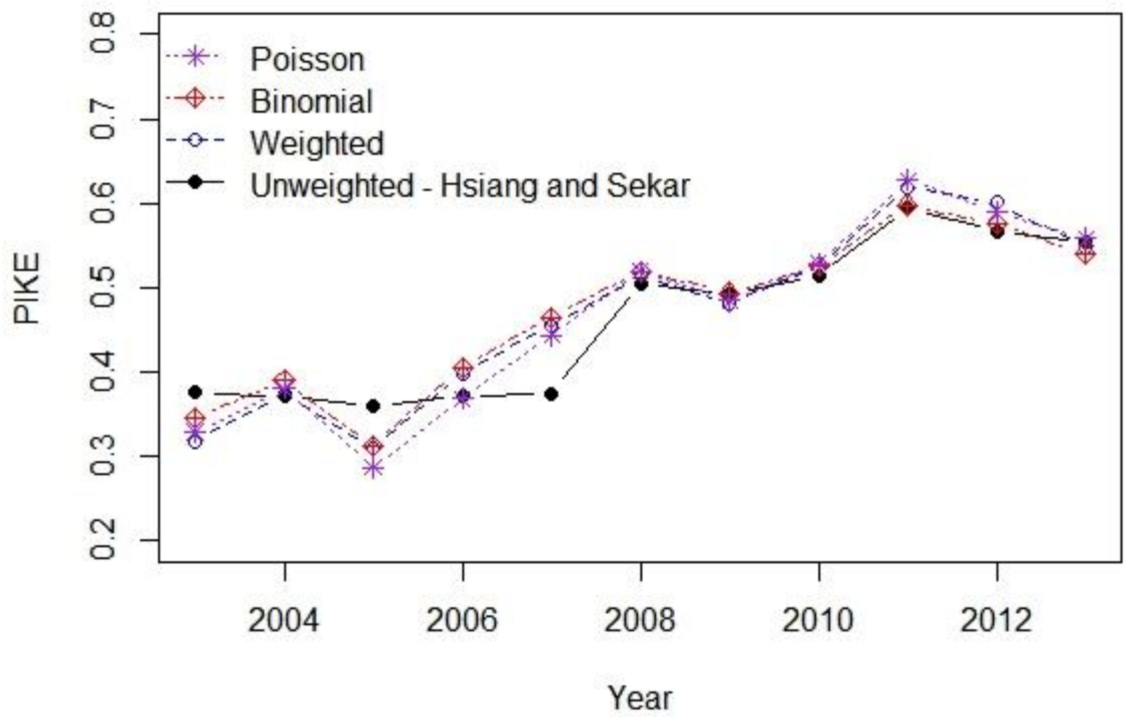

Figure 5: Results of fitting different models to the PIKE data

All three standard approaches that I have outlined for analysing these data show a similar pattern with no discontinuity between 2007 and 2008. They are, as you would expect, also more similar to the aggregate PIKE. The one result that sticks out is that of Hsiang and Sekar who have fundamentally missed the point by ignoring the fact that the total number of carcasses found at a

[4] Their simplest model includes log(number of carcasses not illegally killed +1) as a term in the model. They should use log(total number of carcasses found) and it should be an offset – meaning that it's coefficient is set to 1.

[5] Just a site effect and year effect.

site in any year varies hugely from site to site. If it was the case that the varying number of carcasses did not matter, then you would expect the results to look similar. This is not the case here and Hsiang and Sekar's analysis does not properly take account of the properties of the data.

Summary

The number of carcasses found is not the same in each year or at each site. Any summaries of the data or analyses must account for this. When this is done there is no evidence of the sharp discontinuity between 2007 and 2008 that Hsiang and Sekar claim.

Fiona Underwood

August 2016

Details of Models

Which model	Distribution	Linear predictor
Hsiang and Sekar	$p_{ij} \sim Normal(\mu_{ij}, \sigma^2)$	$\mu_{ij} = \beta_0 + site_i + year_j$
Weighted regression	$p_{ij} \sim Normal(\mu_{ij}, \sigma^2 / w_{ij})$	$\mu_{ij} = \beta_0 + site_i + year_j$ $w_{ij} = n_{ij}$
Binomial	$y_{ij} \sim Binomial(\theta_{ij}, n_{ij})$	$logit(\theta_{ij}) = \beta_0 + site_i + year_j$
Poisson	$y_{ij} \sim Poisson(\lambda_{ij})$	$\log(\lambda_{ij}) = \beta_0 + \log(n_{ij}) + site_i + year_j$
y_{ij} = **number of illegally killed carcasses found at site** i **in year** j n_{ij} = **total number of carcasses found at site** i **in year** j $p_{ij} = y_{ij}/n_{ij}$ $site_i$ = **site effect** $year_j$ = **year effect**		